ESTIMACION DE UNA PROPORCION BINOMIAL MEDIANTE METODOS BAYESIANO

ANDREA JIMENA HERRERA GOMEZ

Trabajo de grado como requisito parcial para optar al título de Profesional en Matemáticas con Énfasis en Estadística

Director
JAIRO ALFONSO CLAVIJO MÉNDEZ
Magister en Ciencias - Estadística

UNIVERSIDAD DEL TOLIMA
FACULTAD DE CIENCIAS
PROGRAMA DE MATEMATICAS Y ESTADISTICA
IBAGUE - TOLIMA
2016

UNIVERSIDAD DEL TOLIMA

FACULTAD DE CIENCIAS

PROGRAMA DE MATEMÁTICAS CON ÉNFASIS EN ESTADÍSTICA

ACTA DE SUSTENTACIÓN TRABAJO DE GRADO

TÍTULO: ESTIMACIÓN DE PROPORCIÓN BINOMIAL MEDIANTE MÉTODOS BAYESIANOS

AUTORES: ANDREA JIMENA HERRERA GOMEZ Cód. 070200112007

DIRECTOR: JAIRO ALFONSO CLAVIJO

JURADOS: ALFONSO SANCHEZ HERNANDEZ
YURI MARCELA GARCIA SAAVEDRA

CALIFICACIÓN: Tres punto ocho (3.8)

X APROBÓ ___ REPROBÓ

OBSERVACIONES: Aprobado

FIRMAS

ALFONSO SANCHEZ HERNANDEZ	YURI MARCELA GARCIA SAAVEDRA
Jurado 1	Jurado 2
JAIRO ALFONSO CLAVIJO	JEAMMY JULIETH SIERRA HERNÁNDEZ
Director del Trabajo	Director del Programa

Ciudad y fecha: Ibagué, 05 de julio de 2016

AGRADECIMIENTOS

A Dios por darme el discernimiento para generar en mí la habilidad de consulta e iniciativa para la consecución de información que me permitió consolidar los conceptos y su aplicabilidad a mi tesis.

A Jairo Alfonso Clavijo, profesor de la facultad de Ciencias de la Universidad del Tolima, asesor y director de este trabajo quien me apoyo dándome su enseñanza, colaboración y dedicación para poder realizar esta tesis.

A mi familia por su colaboración, ayuda y credibilidad en mis habilidades para lograr mis objetivos.

CONTENIDO

RESUMEN

El objetivo de este trabajo es presentar un bosquejo para la estimación bayesiana de parámetros bajo los supuestos de un modelo binomial como se explica con el ejemplo 2, satisfaciendo uno de los objetivos que se plantearon en la obra, como es, proveer los conceptos y elementos fundamentales para la realizar una estimación.

El presente material surge de la evolución del seminario de Estadística Bayesiana que se ha dictado en la Universidad del Tolima durante los últimos semestres como complemento a la introducción de la Estadística. Además, queremos seguir incentivando el uso de métodos computacionalmente basados en la simulación (Monte Carlo con Cadenas de Markov, R, Minitab entre otros), son métodos que permitieron resolver muchos de los problemas analíticos de la estadística bayesiana y la clásica, lo que ha permitido ajustar prácticamente cualquier modelo de probabilidad. De igual forma existe otro programa “OpenBugs” que es un pequeño desarrollo del WinBugs que ha contribuido a la difusión y uso práctico de los métodos Bayesianos.

ABSTRACT

The objective of this work is to present a sketch for the Bayesian estimation of parameters under the assumptions of a binomial model as explained by example 2, satisfying one of the objectives that were raised in the work, namely, to provide the concepts and elements to make an estimate.

The present material arises from the evolution of the Bayesian Statistics seminar that has been given at the University of Tolima during the last semesters as a complement to the introduction of Statistics. In addition, we want to continue to encourage the use of computationally based simulation methods (Monte Carlo with Markov Chains, R, Minitab among others), are methods that solved many of the analytical problems of Bayesian and classical statistics, which has Almost any model of probability can be adjusted. There is also another "OpenBugs" program which is a small development of WinBugs that has contributed to the dissemination and practical use of Bayesian methods.

INTRODUCCIÓN

Entre los numerosos problemas de tipo teórico que se estudian en estadística se encuentra el de estimar proporciones poblacionales, bien sea para dos o para varias categorías. El estudio de las proporciones es un problema viejo que ha sido extensamente estudiado por numerosos investigadores sin que podamos decir que en el momento haya una teoría completa que de sustento a una solución. Prácticamente desde comienzos del siglo pasado se han dado numerosas reglas y procedimientos para lograr estimaciones puntuales y la construcción de intervalos de Confianza. Pearson a principios del siglo pasado propuso métodos que posteriormente fueron revisados por otros autores y que han conducido a distintos métodos de estimación, como se puede constatar en el trabajo de grado presentado por Erika Tinoco, en la Universidad del Tolima. La mayor parte del trabajo de investigación en proporciones está enfocado al caso de proporciones binomiales, es decir, en caso que la población se divida en dos categorías, se desea estimar la proporción que representa una de ellas. A pesar de las numerosas literaturas al respecto, no podemos afirmar que este problema sea un caso terminado y siempre aparecen nuevos aportes teóricos a su solución. Actualmente se hace uso de un amplio método propuesto por Wald utilizando aproximación normal el cual exige tamaños de muestra grandes que deben ser calculados con antelación.

Este trabajo vuelve sobre el tema de la estimación de proporciones pero se hace desde un punto de vista diferente: utiliza el enfoque bayesiano que ha mostrado ser novedoso, efectivo e innovador y genera aportes a la comprensión del problema de las proporciones de gran importancia práctica.

1. HISTORIA DE LA ESTADÍSTICA BAYESIANA

Las bases de la estadística bayesiana se remontan dos siglos atrás, cuando surge el interés por los métodos Bayesianos como aplicación o solución de problemas, ya sea de manera práctica o en término conceptuales donde se da un gran avance para la época. De igual forma, se inicia la etapa evolutiva para el manejo de herramientas computacionales más eficientes y rápidas para la Estadística Bayesiana, permitiendo que los programas y paquetes vinculados sean más eficientes, además, surgen aportes para la solución de problemas como: Teoría de colas, modelos lineales dinámicos entre otros. Asimismo, el estudio de la Estadística Bayesiana se aplica en los siguientes campos, por ejemplo, la Econometria, el Análisis clìnico, Psicometrìa entre otros. Para profundizar(Quintana, 1996).

Entre los 80 y 90 los problemas sonde fácil solución. Sin embargo, muchos problemas dependen del cálculo con un alto porcentaje de problemas de fàcil solución, pero en varios casos dependen del cálculo de funciones de distribuciòn cuyo tratamiento se volvió difícil y engorroso; aparecieron mètodos efectivos de càlculo numérico que facilitan la obtención de valores simulados. Se aprovecharon contribuciones de distintos procedimientos de integraciòn numérica y aproximaciones de tipo Legendre y Laplace. La idea de muestreo y remuestreo dieron pautas para el manejo del algoritmo Gibbs Samplery otros, con el mismo principio (Metrópolis y Metrópolis – Hastings en particular).

El método que tomó fuerza fue el Método Montecarlo basado en Cadenas de Markov también conocido como: MCMC (Markov Chain Monte Carlo), una herramienta que permite resolver problemas de gran complejidad (Gilks W. R., 1996)

1.1 NUEVAS HERRAMIENTAS DE TRABAJO

Se podría decir que el análisis bayesiano consiste en muestrear por simulación, y usar simultáneas distribuciones de probabilidad.

El manejo de algunas distribuciones puede hacerse analíticamente en situaciones simples, pero en algunos casos es necesario implementar el uso computacional para muestrear y simular.

2. PRELIMINARES

2.1 OBSERVACIÓN.

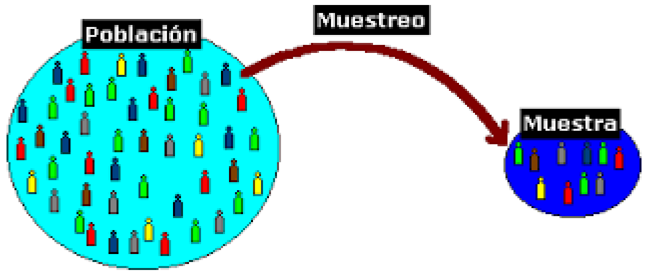
Las variables aleatorias pueden ser clasificadas en numéricas y categóricas. En este trabajo se manejaran solamente variables aleatorias categóricas para las cuales la distribución de probabilidad se define así:

$$P(X = x) = \begin{cases} P_i & si \quad x = x_i \ (i = 1,2, \dots, k) \\ 0 & en \quad otro\ caso \end{cases}$$

K es el número de categorías de la variable X.

El caso $K = 2$ corresponde a una población con dos clases de elementos a los que llamaremos éxitos y fracasos.
En principio se puede suponer que la población es finita con N elementos y que en ella hay A Éxitos y $N - A$ fracasos.
La fracción $\pi = \frac{A}{N}$ se llama proporción de éxitos en la población. En la práctica $\boldsymbol{\pi}$ no es conocido y se busca estimar a partir de una muestra; por esta razón, el problema se conoce como “estimación de la proporción” la cual se generaliza en caso que la población tenga más categorías de elementos.

Figura 1. Estimación de la proporción



Fuente: Presentación Name on emaze

Si hay K categorias cada una de ellas representa un subconjunto C_i de la población. Y sea A_i el número de elementos en C_i entonces $\pi_i = \frac{A_i}{N}$ será la proporción correspondiente a la categoría C_i. Esto lleva a otro problema, estimar π_i donde $i = 1,2, \dots, k$.

El primer problema se conoce como estimación de una proporción binomial, el segundo como estimación de proporciones multinomiales.
Este trabajo se enfoca en la estimación de una proporción binomial usando técnicas bayesianas en vez de aplicar la forma tradicional Frecuentista.

El problema está asociado con la distribución binomial por la siguiente razón:
Al elegir aleatoriamente un elemento de la población y que dicho elemento sea un éxito, tiene probabilidad $\pi = \frac{A}{N}$, por tanto, la probabilidad de que ese elemento no sea éxito es$1 - \pi = 1 - \frac{A}{N} = \frac{N}{N} - \frac{A}{N} = \frac{N-A}{N}$,cociente entre el total de no éxitos y el tamaño de la población, por tanto:

$$P(X = \text{éxito}) = \pi$$

$$P(X = \text{fracaso}) = 1 - \pi$$

Para estudiar la v.a. X usaremos una variable auxiliar Y, definida así:

$$Y = \begin{cases} 1 \;\; si\ X\ es\ \text{éxito} \\ 0 \;\; si\ X\ es\ \text{fracaso} \end{cases}$$

Así $\mathbf{Y}$ es una variable aleatoria con distribución Bernoulli, es decir, $\mathbf{Y} \sim \mathbf{b}(\boldsymbol{\pi})$. En consecuencia:

$$\mathbf{E}(\mathbf{Y}) = \boldsymbol{\pi}$$
$$\mathbf{V}(\mathbf{Y}) = \boldsymbol{\pi}(\mathbf{1} - \boldsymbol{\pi})$$

Si se toma una muestra aleatoria de $\boldsymbol{n}$ elementos de la población con reemplazamiento, digamos: $Y_1, \dots \dots, Y_n$ con $Y_i \sim^{id} Y$ y si llamamos

$$\mathbf{S} = \boldsymbol{Y_1} + \boldsymbol{Y_2} + \dots \dots . \dots + \boldsymbol{Y_n}$$

Se tiene que $\mathbf{S}$ representa el número de éxitos en esa muestra y $\mathbf{S} \sim \mathbf{B}(\mathbf{n}, \boldsymbol{\pi})$, y por tanto,

$$E(\mathbf{S}) = E\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} E(Y_i)\, \pi.$$

$$\mathrm{V}(\mathbf{S}) = \mathrm{V}\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} V(Y_i) + \sum_{i<j} cov\left(Y_i, Y_j\right)$$

Donde $2\sum_{i<j} cov\left(Y_i, Y_j\right)$ vale cero a causa de la independencia, por lo tanto

$$V(S) = \sum_{i=1}^{n} V(Y_i) = n\pi(1 - \pi)$$

Se define:

$$p = \frac{S}{n}$$

Se tiene

$$E(p) = E\left(\frac{S}{n}\right) = \frac{1}{n} E(S) = \frac{n\pi}{n} = \pi$$

Lo que permite utilizar a $\boldsymbol{p}$ como estimador insesgado de π.

Tabla 1. Nomenclatura para probabilidad y proporción

Población	Muestra
Proporción	Estimador
$\boldsymbol{\pi}$	p

Nota: $\boldsymbol{p}$ Es un estimador insesgado de $\boldsymbol{\pi} = \boldsymbol{E}(\boldsymbol{P}) = \boldsymbol{\pi}$. Y es convergente en probabilidad, ya que su varianza disminuye al aumentar el tamaño muestral, ¿Qué es convergente en probabilidad.

$$V(P) = \pi \frac{(1-\pi)}{n}.$$

La forma corriente de estimar utilizada es la aproximación normal propuesta por WALD, según la cual:

$$\frac{S \sim E(S)}{\sqrt{V(S)}} \simeq N(0,1)$$

Pero ésta no es la mejor manera de estimar (Sachs, 2013), $V(S)$depende del tipo de muestreo utilizado, más exactamente si se hace con repetición o no.

Si se hace con repetición, la probabilidad de elección de cada elemento es $\boldsymbol{\pi} = \frac{\boldsymbol{A}}{\boldsymbol{N}}$ ya que el tamaño de la población no cambia, esto corresponde a una distribución binomial. Si se hacen SIN repetición, la probabilidad cambia después de cada extracción, y la probabilidad se ajusta a un modelo hipergeometrico lo que lleva a:

$$V(S) = \begin{cases} n\pi(1-\pi) \; si\ es\ una\ M.A.S\ con\ \text{reposición} \\ n\,\frac{A}{N}\left(1-\frac{A}{N}\right)\left(1-\frac{n}{N}\right) \; si\ se\ usa\ M.A.S\ \ sin\ \text{reposición} \end{cases}$$

Y como $\frac{\boldsymbol{A}}{\boldsymbol{N}} = \boldsymbol{\pi}$ se tiene

$$V(S) = \begin{cases} n\pi(1-\pi) \; si\ se\ usa\ \ M.A.S\ con\ \text{reposición} \\ \boldsymbol{n\,.\,\pi(1-\ \pi)}\left(\boldsymbol{1-\frac{n}{N}}\right) \; si\ se\ usa\ M.A.S\ \ sin\ \text{reposición} \end{cases}$$

Si la población es infinita o muy grande la extracción de una unidad no afecta su tamaño y entonces el M.A.S CON y el M.A.S SIN son prácticamente iguales, y en tal caso

$$\boldsymbol{V(S) = n\pi(1-\pi)}$$

El caso problemático ocurre cuando la población es pequeña y se usa muestreo sin reemplazamiento.

Nos concentraremos en el caso binomial, es decir, poblaciones muy grandes o muestreo aleatorio simple CON reemplazamiento.

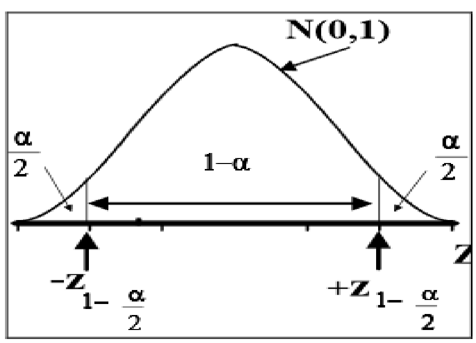
Según la propuesta de WALD de aproximación normal, si $\boldsymbol{n}$ es suficientemente grande, se tiene:

$$\frac{S-n\pi}{\sqrt{V(S)}} = \frac{S-n\pi}{\sqrt{n\pi(1-\pi)}} \simeq N(0,1)$$

$$\Rightarrow Z = \frac{S-n\pi}{\sqrt{\frac{n\pi(1-\pi)}{n^2}}} = \frac{p-\pi}{\sqrt{\frac{\pi(1-p)}{n}}} \simeq N(0,1)$$

Definición 1. Un intervalo de confianza es un intervalo que tiene a lo menos un extremo aleatorio y es construido de manera tal que el parámetro de interés que se estima está contenido en dicho intervalo con una probabilidad $1-\alpha$, llamada coeficiente de confianza.

Figura 2. INTERVALO DE CONFIANZA POR APROXIMACIÒN NORMAL



Fuente: Mario Orlando Suárez Ibujes

Por tanto, la construcción del IC se basa en el siguiente procedimiento:

$$Pr\left(-Z\alpha/_2 \leq \frac{\hat{p}-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}.\sqrt{1-\frac{n}{N}}} \leq Z\alpha/_2 \leq \frac{\hat{p}-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}.\sqrt{1-\frac{n}{N}}}}\right) = 1-\alpha \qquad [1]$$

De donde se deduce:

$$Pr\left(-Z\alpha/_2\sqrt{\frac{\pi(1-\pi)}{n}}.\sqrt{1-\frac{n}{N}} \leq \hat{p}-\pi \leq Z\alpha/_2\sqrt{\frac{\pi(1-\pi)}{n}}.\sqrt{1-\frac{n}{N}}\right)[2]$$

$$= Pr\left(\hat{p}-Z\alpha/_2\sqrt{\frac{\pi(1-\pi)}{n}}.\sqrt{1-\frac{n}{N}} \leq \pi \leq \hat{p}+Z\alpha/_2\sqrt{\frac{\pi(1-\pi)}{n}}.\sqrt{1-\frac{n}{N}}\right) = \qquad [3]$$

Luego IC para π es

$$\hat{p} \pm \left(Z\alpha/_2\sqrt{\frac{\pi(1-\pi)}{n}}.\sqrt{1-\frac{n}{N}}\right)[4]$$

En la práctica del cálculo del IC se cambia π (desconocido) por $\hat{p}$ lo que obliga a un pequeño cambio (Cochran, sampling techniques)

$$\hat{p} \pm \left(Z\alpha/_2\sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}.\sqrt{1-\frac{n}{N}}\right)$$

Pero se usa n en vez de $n-1$ con lo cual el IC es:

$$\hat{p} \pm Z\alpha/_2\sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}\ .\sqrt{1-\frac{n}{N}} \text{ (Wald)}$$

El caso problemático, mediante razonamientos parecidos lleva a la significante expresión para el IC:

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}\cdot\sqrt{1-\frac{n}{N}}$$

Factor de corrección por finitud.

El cálculo del tamaño de muestra se podría resumir en dos pasos así:

$n_\infty = \frac{z^2pq}{\varepsilon^2}$ (Aquí se debe discutir como determinar p y q)

$n = \frac{n_\infty + 1}{1+\frac{n_\infty}{N}}$ (Esta corrección se usa cuando la población es finita, de tamaño N)

Existen métodos exactos para estimar la proporción a la manera clásica. Entre ellos se puede mencionar.

1. Score de Wilson(1927)
2. Coppler - Pearson (1934)
3. Blyth - Hutchinson (1964)
4. Agresti & Coull (1998)

Pero se tiene como ventaja el hecho de no depender de la distribución de estimadores.

5. El método no paramétrico conocido como Bootstrap.

Tabla 2. Métodos exactos

AUTOR	MÉTODO
Score de Wilson (1927)	intervalo de puntuación de Wilson
Coppler – Pearson (1934)	Intervalo “Exacto”
Blyth _ Hutchinson (1964)	Estimación de los parámetros de mezclas de distribuciones binomiales

3. CONCEPTOS BÁSICOS

3.1 TEOREMA DE BAYES

Con el Teorema de Bayes, surge la Teoría bayesiana permitiendo tomar como base el cálculo de probabilidad de un evento apoyado en un conocimiento a priori. (Bayes T. , 1763).

El Teorema permite calcular la probabilidad posterior a partir de la verosimilitud y probabilidad anterior de los datos.

Para enfocarnos en este Teorema se toman algunos conceptos:

- La probabilidad condicional es establecida por la aplicación de una norma de distribución de probabilidad según el carácter de dicha distribución.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \; ; P(B) > 0$$

De aquí

$$P(A \cap B) = P(A|B)\, P(B)$$

De igual forma

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

De tal manera que

$$P(A \cap B) = P(A|B)\, P(A)$$

Donde especifica que:

$$P(A|B) = \; P(B|A)$$

“La probabilidad de A dado B = a la probabilidad de B dado A”

Ejemplo 1:

El 20% de las profesoras del colegio Técnico Alemán son profesionales y otro 20% son Técnicos. El 75% de las profesionales ocupan un puesto directivo y Académico y el 50% de los Técnicos también, mientras que los no profesionales y los no Técnicos solamente el 20% ocupa un puesto académico. ¿Cuál es la probabilidad de que una de las profesoras elegido al azar sea profesional?

0.2	Profesionales	0.75
0.2	Técnicos	0.5
0.6	Otros	0.2

Utilizando la formula Bayesiana será:

$$P(Profesinales|\text{Técnicos}) = \frac{(0.2)(0.75)}{(0.2)(0.75) + (0.2)(0.5) + (0.6)(0.2)} = 0{,}40$$

Se puede generalizar

Sea $B_1, B_2, \dots, B_n$ son sucesos mutuamente excluyentes, se tiene que para cualquier evento A:

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

Este es el resultado correspondiente a la probabilidad total.

De esta manera se genera el Teorema de Bayes:

Sea $\{A_1, A_2, \dots, A_i, \dots, A_n\}$ un conjunto completo de sucesos. Entonces B un suceso cualquiera del que se conocen las probabilidades condicionales $P(B|A_i)$. Luego, la probabilidad de $P(A_i|B)$se da por la expresión:

$$P(A_i|B) = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

Dónde:

- $P(A_i)$ son las probabilidades a priori.

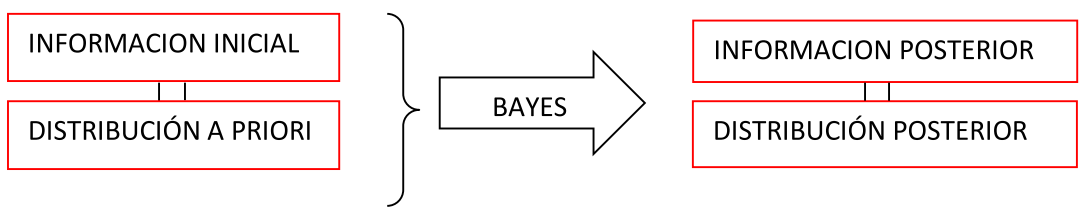
- $P(B|A_k)$Probabilidad de las hipótesis de A_i

- $P(A_i|B)$: probabilidad posterior.

El uso de información conocida a priori es la parte destacada de la inferencia bayesiana (LÓPEZ DE CASTILLA, 2011)

Figura 3. Ilustración del Teorema de Bayes

Distribución posterior $\propto$ función de verosimilitud $\times$ distribución a priori.



Fuente: autor

3.2 FUNDAMENTOS TEÓRICOS

Se usará el concepto de inferencia para estimar parámetros dentro de un espacio muestral y modelar el parámetro de su interés, para contrastar hipótesis y tomar decisiones. Dentro de la inferencia se aplica ***la filosofía bayesiana*** que enfrenta los conceptos de *inferencia clásica y Bayesiana: en* la ***inferencia clásica (ò Frecuentista)*** el parámetro θ es tratado como una constante, en ***la Bayesiana*** es considerada como una variable aleatoria, estos conceptos no coinciden, de tal manera, las características propias de cada aplicación determina cual es la adecuada para utilizar.

Cualquier problema estadístico evaluado bajo el método bayesiano implica obtener y utilizar la familia paramétrica adecuada, por tanto, se tiene en cuenta el parámetro θ.

A raíz de la aparición de la Estadística Bayesiana se han desarrollado técnicas de estimación en lo que se puede incorporar el conocimiento previo y subjetivo que se toma del problema, lo que abre las puertas a un campo amplio y bastante inexplorado de la estadística.

Este es particularmente el punto de interés de este trabajo: es utilizar métodos bayesianos para estimar una proporción binomial.

3.2.1 Diferencias entre la Inferencia Frecuentista y Bayesiana. Si se toma un parámetro poblacional θ sobre el cual se realizaran inferencias y se obtendrá un modelo de probabilidad $\boldsymbol{f}(\boldsymbol{y}|\boldsymbol{\theta})$ donde establece la probabilidad de los datos observados $\boldsymbol{\mathcal{Y}}$ bajo diferentes valores de $\boldsymbol{\theta}$.

La diferencia fundamental entre la teoría frecuentista y la bayesiana es:

- La inferencia frecuentista trata de estimar un parámetro $\boldsymbol{\theta}$, que se considera fijo (constante) pero desconocido, la inferencia Bayesiana trata de estimar un parámetro $\boldsymbol{\theta}$, que se considera una variable aleatoria a la cual se le asocia a una distribución de probabilidad $\boldsymbol{f}(\boldsymbol{\theta})$ (a priori).

- La inferencia frecuentista no tiene como finalidad producir una conclusión dicotómica (significación o no significación, rechazo o aceptación, etc.), la inferencia Bayesiana se basa en $\boldsymbol{f}(\boldsymbol{\theta}|\boldsymbol{y})$ conocida como la distribución posterior para el parámetro luego de observados los datos.

- La Estadística Bayesiana calcula el grado de credibilidad y se supone comenzando con una probabilidad a priori y actualizándola en la presencia de evidencia por medio del teorema de Bayes.

(Correa, 2005) Crea una diferenciación entre la estadística clásica y la bayesiana, explicada en la siguiente tabla.

Tabla 3. DIFERENCIA ENTRE LA ESTADÍSTICA FRECUENTISTA Y LA BAYESIANA

CARACTERÍSTICAS	FRECUENTISTA (CONSTANTE)	BAYESIANA (VARIABLE ALEATORIA)
PARÁMETRO DE INTERÉS($\boldsymbol{\theta}$)	CONSTANTE DESCONOCIDA	VARIABLE ALEATORIA
DISTRIBUCIÓN A PRIORI	NO EXISTE	EXISTE Y ES EXPLICITA
MODELO MUESTRAL	SE ASUME	SE ASUME
DISTRIBUCIÓN POSTERIOR	NO EXISTE	EXISTE Y DERIVA
RAZONAMIENTO	INDUCTIVO	DEDUCTIVO

3.3 FILOSOFÍA BAYESIANA

En la filosofía bayesiana la incertidumbre se puede realizar mediante el cálculo de probabilidades, así como se explica en los siguientes argumentos:

- La filosofía frecuentista evalúa qué tan verosímiles son los datos de acuerdo con diferentes valores hipotéticos para el parámetro desconocido $\boldsymbol{\theta}$. Las afirmaciones acerca de la probabilidad de observar los datos que se tiene, dados diferentes valores hipotéticos para el parámetro, se resumen en un Intervalo de Confianza.

- La filosofía Bayesiana evalúa qué tan verosímiles son los valores del parámetro $\boldsymbol{\theta}$, dados los datos observados. Por tanto, la afirmación a la que se llega se refieren a la probabilidad de que la cantidad desconocida tome un determinado valor en un cierto Intervalo de Credibilidad.

- Se suministra una especificación justificada de $f(\theta)$ (es una distribución de probabilidad sobre la distribución inicial Θ) donde el método bayesiano proporciona el medio con que se realizará una estimación entre el espacio θ. Pueden ser consultados en cualquier texto básico (Bolstad, 2004).

Además, como $f(\theta)$ es la información a priori y $f(\theta|Y)$ es la información posterior que se obtiene de la muestra. Se podría tener en cuenta las dos condiciones para modelar nuestra incertidumbre sobre θ.

- "A priori" de una cantidad $\boldsymbol{p}$ desconocida, es la distribución de probabilidad que enuncia alguna incertidumbre acerca de $\boldsymbol{p}$ antes de tomar en cuenta los "datos". Aplicando el Teorema de Bayes se podrá obtener estimaciones de los parámetros de la distribución posterior en base a las distribuciones a priori (véase en el anexo A).

- "Posterior" es igual a la probabilidad *a priori* por la información muestral. Atendiendo a la propiedad de las familias conjugadas , la combinación de una distribución *a priori* beta y una verosimilitud con distribución binomial da como resultado una distribución *a posteriori* del tipo beta con parámetros $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ (véase en el anexo B).

3.4 INFERENCIA BAYESIANA EN UN CASO BINOMIAL.

La distribución a priori juega un papel fundamental en el análisis bayesiano, ya que calcula el grado de conocimiento inicial que posee los parámetros en estudio. Si tenemos en cuenta que el parámetro es desconocido, según algunos autores piensan que debe existir un conocimiento a priori $\boldsymbol{f(\theta)}$, establecido por precedentes experiencias donde la información será incorporada en el análisis, a partir de un modelo adecuado de $\boldsymbol{f(\theta)}$, para contribuir de cierta forma a una mejor definición de $\boldsymbol{f(\theta/y)}$**,** y a su vez, que no influya dentro de los posibles valores del parámetro**.**

Según el Teorema Bayes, al multiplicar la probabilidad a priori por la verosimilitud, se logra la distribución de probabilidad posterior, siendo esta la probabilidad de la distribución condicional determinada por los datos.

En la distribución a priori, los parámetros son llamados hiperparametros, para diferenciarlos de los parámetros del modelo. Por ejemplo, utilizando una distribución Beta para modelar la distribución del parámetro $\boldsymbol{\theta}$, hay que tener en cuenta:

- θ es una variable aleatoria con distribución Beta

- α y β son parámetros de la distribución a priori (distribución beta), y por lo tanto son hiperparametros

3.5 DISTRIBUCIÓN A PRIORI NO INFORMATIVA

En la distribución **"A priori"** puede suceder que el conocimiento previo de $\boldsymbol{\theta}$, sea amplio y bien determinado, además de que puede influir fuertemente en la distribución posterior. Si esto sucede, se dice que la distribución **A priori** de $\boldsymbol{\theta}$ es informativa, o por el contrario si el conocimiento sobre $\boldsymbol{\theta}$ es vago y no hay razones para creer que la influencia de dicho conocimiento es fuerte en la distribución posterior, se dice que la distribución a priori es no informativa.
Una distribución a priori parece ser más o menos informativa de acuerdo con los valores de sus parámetros, los cuales deben ser seleccionados adecuadamente según muestra de interés de que la A priori sea $\boldsymbol{\theta}$no informativa o Difusa. Además, dentro de las distribuciones a priori no informativas se tiene el siguiente criterio.

El autor (Jeffreys, 1961) planteó una clase de distribuciones a priori no informativas para el caso de espacios paramétricos infinitos. La construcción de esta clase se basa en buscar simultáneamente invariancia ante transformaciones y proveer la menor información a priori en relación a la información muestral. Mayor información se encuentra en (Casella, 2001) y (Lindley, 1962).

Jeffreys implantó una distribución a priori, con la propiedad de invariancia.

$$f(\theta) \propto \sqrt{I(\theta)}$$

Donde$I(\theta)$ es la matriz de información de Fisher, por otro lado

$$E_\theta\left[\frac{d^2 \log(f(y/\theta))}{d\theta^2}\right]$$

Si $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_k \end{pmatrix}$ es un vector, entonces

$$f(\theta) \propto \sqrt{det(I(\theta))}$$

El i – j ésimo elemento de esta matriz es:

$$I_{ij} = E_\theta \left[\frac{d^2 \log(f(y/\theta))}{d\theta_i d\theta_j} \right]$$

O también

$$I(\theta) = Var_X \left[\frac{d}{d\theta} \log f(y/\theta) \right] = E_X \left[\left(\frac{d}{d\theta} \log f(y/\theta) \right)^2 \right] = E_X \left[\frac{d^2}{d\theta^2} \log f(y/\theta) \right]$$

Es la esperanza de Fisher.

Además, se ha implementado el uso de distribuciones a priori no informativa, ya que ha mostrado propiedades frecuentistas (Canavos, 1998).

Como ejemplo de distribución a priori No informativa se tiene Beta (1,1)

$$\boldsymbol{IC_{no\ informativa}} = [\boldsymbol{B(x+1, n-x+1, \alpha/2)}, \boldsymbol{B(x+1, n-x+1, 1-\alpha/2)}]$$

Con $f(\pi)$= constante, para todos los valores de π.

$$f(\pi) = \frac{\Gamma(2)}{\Gamma(1)\Gamma(1)} \pi^{(1-1)} (1-\pi)^{(1-1)}$$

Que es una distribución uniforme sobre el intervalo $[0,1]$.

Otra distribución No informativa es $B(1/2, 1/2)$ explicada en el ejemplo 3.

Después de observar $\boldsymbol{x}$ éxitos en $\boldsymbol{n}$ pruebas, la ***distribución posterior*** de $\boldsymbol{p}$ es una distribución Beta con parámetros *(x* + 1/2, *n* - *x* + 1/2).

$$\boldsymbol{IC_{Jeffreys}} = [\boldsymbol{B(x+1/2, n-x+1/2, \alpha/2)}, \boldsymbol{B(x+1/2, n-x+1/2, 1-\alpha/2)}]^{*}$$

Hace referencia que tanto el intervalo de Wilson como el intervalo de jeffreys funcionan perfectamente, además, la cobertura de Jeffreys es muy cercana al nivel nominal, aunque los valores sean suficientemente pequeños de $\boldsymbol{n}$ (Lawrence D. Brown, 2001).

Ejemplo 3:

Supóngase que $Y \sim B(n, \theta)$ y se quiere hallar una distribución a priori no informativa para $\boldsymbol{\theta}$.

$$f(y/\theta) = P(y/\theta) = \binom{n}{y} \theta^{y}(1-\theta)^{n-y}$$

$$L_n\big(f(y/\theta)\big) = L_n\left(\binom{n}{y}\right) + yL_n(\theta) + (n-y)L_n(1-\theta)$$

$$\frac{dL_n\big(f(y/\theta)\big)}{d\theta} = \frac{y}{\theta} + \frac{n-y}{1-\theta}$$

$$\frac{d^2L_n\big(f(y/\theta)\big)}{d\theta^2} = \frac{y}{\theta^2} + \frac{n-y}{(1-\theta)^2}$$

De aquí:

$$E\left[\frac{d^2L_n\big(f(y/\theta)\big)}{d\theta^2}\right] = E\left[-\frac{y}{\theta^2} + \frac{n-y}{(1-\theta)^2}\right] = \left[-\frac{n\theta}{\theta^2} + \frac{E(n-y)}{(1-\theta)^2}\right] = \frac{n}{\theta(1-\theta)}$$

O sea: $I(\theta) = \left(\frac{\sqrt{n}}{\sqrt{\theta}\sqrt{1-\theta}}\right)^2$

En consecuencia $P(\theta) = \sqrt{I(\theta)} = \frac{\sqrt{n}}{\sqrt{\theta}\sqrt{1-\theta}} = \sqrt{n}\sqrt{\theta}^{-1}\sqrt{1-\theta}^{-1}$

O sea $P(\theta) \propto \theta^{\frac{1}{2}-1}(1-\theta)^{\frac{1}{2}-1}$

Que es una $Beta\left(\frac{1}{2}, \frac{1}{2}\right)$.

*Es decir, que cuando se tiene una muestra proveniente de una distribución binomial de parámetro θ (pruebas de éxitos), la distribución a priori no informativa, para θ es una Beta de parámetros $\frac{1}{2}, \frac{1}{2}$.

3.6 DISTRIBUCIÓN A PRIORI INFORMATIVA

Cuando se habla de distribución informativa, se deduce que la información recogida del modelo probabilístico a priori, permiten determinar el comportamiento a posteriori del parámetro. Además, se crea una familia de distribuciones como información a priori del parámetro, el cual genera información fundada en su comportamiento. Es así como las distribuciones conjugadas logran establecerse como informativas.

3.7 DISTRIBUCIÓN POSTERIOR

la distribución posterior o distribución final se obtiene al aplicar el teorema de bayes a la distribución a priori de la variable que se estableció; permitiendo que haya una mezcla entre la información previa a la “experiencia” con la información objetiva de los datos.

Además, con esta distribución se puede calcular el estimador puntual bayesiano, las probabilidades de hipótesis creadas con opción de dos o más hipótesis y los intervalos de alta densidad relacionada al parámetro que se desea estimar. La distribución posteriori se tendrá en cuenta dentro de la inferencia bayesiana donde se encontraran los parámetros o cantidades desconocidas de los modelos, por ejemplo el valor esperado posterior, son resultados asintóticos para solucionar problemas de integración para así obtener resultados.

Sin embargo, la distribución posterior es el estimador bayesiano de $\boldsymbol{\theta}$, y expresa la nueva distribución $\boldsymbol{\theta}$ variada en función de los datos observados.

Cuando no es posible realizar cálculos directos de la distribución a posteriori, se opta por aplicar la metodología MCMC, de igual forma, si esta no tiene una distribución conocida una de las soluciones es crear una muestra con parámetros de

interés pensando en un procedimiento MCMC, donde se simula una cadena de Markow con distribución estacionaria dada por la distribución a posteriori $\boldsymbol{f}(\boldsymbol{\theta}|\boldsymbol{y})$(gilks w. r., 1996).

El ejemplo siguiente reviste gran importancia dentro de este trabajo ya que él satisface uno de los objetivos que se plantearon en la obra, como es, proveer los conceptos y elementos fundamentales para la estimación de una proporción binomial.

Ejemplo 2:

Supongase que el parámetro$p$ de una binomial tienedistribución a priori de la forma:

$$\Phi(p) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}$$

Donde $B(\alpha,\beta)$ es la función Beta

Y que se han observado n ensayos independientes en los que hay x éxitos.

La distribución conjunta de p y x está dada por:

$$f(p|x) = \binom{n}{x} p^x q^{n-x} \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}$$

$$= f(x|p)\Phi(p) \text{ (verosimilitud)}$$

$$f(x) = \int_0^1 \frac{\binom{n}{x} p^x q^{n-x} p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}\, dp$$

$$= \frac{\binom{n}{x}}{B(\alpha,\beta)}\, B \int_0^1 p^{x+\alpha-1}\,(1-p)^{n-x+\beta-1}\, dp$$

$$= \frac{\binom{n}{x}}{B(\alpha,\beta)}\, B(x+\alpha, n-x+\beta)$$

El estimador puntual posterior de p sera:

$$\int_0^1 f(p|x)dp = \int_0^1 \frac{p^{x+\alpha}(1-p)^{n-x+\beta-1}}{B(x+\alpha, n-x+\beta)}\, dp$$

Dando como resultado

$$\frac{B(x+\alpha+1, n-x+\beta)}{B(x+\alpha, n-x+\beta)}$$

Es decir

$$\hat{p}_B = \frac{x+\alpha}{n+\alpha+\beta}$$

Hemos propuesto con este ejercicio que el parametro p tiene distribucion “a priori” beta de parametros α y β, tal como se sigue en el ejercicio 4.

3.7.1 Estimador Posterior

Sea $\{\boldsymbol{y_1}, \boldsymbol{y_2}, \dots, \boldsymbol{y_n}\}$ una muestra aleatoria de $\boldsymbol{f}(\boldsymbol{y}|\boldsymbol{\theta})$ donde $\boldsymbol{\theta}$ es un valor de la v.a. $\boldsymbol{\theta}$ con función de densidad $\boldsymbol{g_\theta}()$.

El estimador posterior de $\mathfrak{S}(\boldsymbol{\theta})$ con respecto a la priori $\boldsymbol{g_\theta}(.)$ está dado por:

$$\mathbf{E}(\mathfrak{S}(\boldsymbol{\theta})|\boldsymbol{y_1}, \boldsymbol{y_2}, \dots, \boldsymbol{y_n})$$

Ejemplo 4:

Suponemos que $\{\boldsymbol{y_1}, \boldsymbol{y_2}, \dots, \boldsymbol{y_n}\}$ es una muestra aleatoria con distribución $\boldsymbol{f}(\boldsymbol{y}|\boldsymbol{\theta}) = \boldsymbol{\theta}^{\boldsymbol{y}}(\mathbf{1}-\boldsymbol{\theta})^{\mathbf{1}-\boldsymbol{y}}$ para $\boldsymbol{y} = \mathbf{0}$ $\boldsymbol{o}$ $\boldsymbol{y} = \mathbf{1}$ y que $\boldsymbol{g_\theta}(\boldsymbol{\theta}) = \cup_{(\mathbf{0},\mathbf{1})}$.

Encontrar estimadores de $\boldsymbol{\theta}$ Y $\boldsymbol{\theta}(\mathbf{1}-\boldsymbol{\theta})$.

$$\boldsymbol{f}(\boldsymbol{\theta}|\boldsymbol{y_1}, \boldsymbol{y_2}, \dots, \boldsymbol{y_n}) = \frac{\boldsymbol{g_\theta}(\boldsymbol{\theta}) \prod_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{f}(\boldsymbol{y_i}|\boldsymbol{\theta})}{\int_{\mathbf{0}}^{\mathbf{1}} \boldsymbol{g_\theta}(\boldsymbol{\theta}) \prod_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{f}(\boldsymbol{y_i}|\boldsymbol{\theta})\, .\, \boldsymbol{d\theta}}$$

$$= \frac{\theta \sum_{i=1}^{n} y_i (1-\theta)^{n-\sum_{i=1}^{n} y_i} I_{(0,1)}(\theta)}{\int_0^1 \theta \sum_{i=1}^{n} y_i (1-\theta)^{n-\sum_{i=1}^{n} y_i}\ d\theta}$$

Al tomar esperanzas, se tiene:

$$\mathrm{E}(\theta | y_1, y_2, \dots, y_n) = \frac{\int_0^1 \theta \sum_{i=1}^{n} y_i (1-\theta)^{n-\sum_{i=1}^{n} y_i}\ d\theta}{\int_0^1 \theta \sum_{i=1}^{n} y_i (1-\theta)^{n-\sum_{i=1}^{n} y_i}\ d\theta}$$

$$= \frac{B(\sum_{i=1}^{n} y_i + 2, n - \sum_{i=1}^{n} y_i + 1)}{B(\sum_{i=1}^{n} y_i + 1, n - \sum_{i=1}^{n} y_i + 1)}$$

$$= \frac{1}{2} \sum_{i=1}^{n} y_i + 1$$

Y en efecto el estimador posterior de bayes para θ es:

$$\frac{1}{n+2}\left(\sum_{i=1}^{n} y_i + 1\right)$$

Que es sesgado, mientras que el estimador mv de θ es $\frac{1}{n}\sum_{i=1}^{n} y_i$ que es insesgado.

Para $\theta(1-\theta)$ se tiene:

$$\mathrm{E}(\theta\ (1-\theta) | y_1, y_2, \dots, y_n) = \frac{\int_0^1 \theta (1-\theta)^{\sum_{i=1}^{n} y_i} (1-\theta)^{n-\sum_{i=1}^{n} y_i} d\theta}{\int_0^1 \theta^{\sum_{i=1}^{n} y_i} (1-\theta)^{n-\sum_{i=1}^{n} y_i} d\theta}$$

$$= \frac{\Gamma(\sum_{i=1}^{n} y_i + 2)\Gamma(n - \sum_{i=1}^{n} y_i + 2)}{\Gamma(n+4)}$$

$$\frac{\Gamma(\mathrm{n}+2)}{\Gamma\ (\sum_{i=1}^{n} y_i + 1)\Gamma(n - \sum_{i=1}^{n} y_i + 1)}$$

$$\frac{(\sum_{i=1}^{n} y_i + 1)(n - \sum_{i=1}^{n} y_i + 1)}{(n+3)(n+2)}$$

3.8 FAMILIAS CONJUGADAS

Definición. Sea $\mathcal{P}:\{p(y|\theta)\ :\ \theta \epsilon \Theta\}$ una familia paramétrica. Una colección de distribuciones de probabilidad $\mathcal{F}$ es una familia conjugada para $\mathcal{P}$ si para todo $p(y|\theta)\ \epsilon\ \mathcal{P}$ y $p(\theta)\epsilon\, \mathcal{F}$ se cumple que $p(y|\theta)\ \epsilon\, \mathcal{F}$.Que es una distribución del mismo tipo que la de $P(\theta)$.

Las distribuciones a priori Beta forma una familia conjugada para la clase de funciones de densidad binomial "La familia Beta es conjugada de la binomial"

Se dice que cuando la distribución a posteriori es de la misma familia que la distribución a priori, es una familia de distribuciones conjugadas.

Nota. Una clase $\wp$ de distribuciones a priori es **conjugada** de una familia $\mathcal{F}$ de función de densidad si cada vez que $P(\theta)\epsilon\, \wp$ y $f(y/\theta)\epsilon\, \mathcal{F}$ se cumple $f(\theta/y)\epsilon\, \wp$.
Al igual que el ejemplo 2 el resto del capítulo es esencial en la estimación de proporción.

Ejemplo 5:
Se considera un parámetro θ que a priori tiene distribución Beta de parámetros α y β y una variable aleatoria Y con distribución de probabilidad binomial de parámetros m y θ.
Se tiene

$$P(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}I_{(0,1)}$$

$$P(y/\theta) = \binom{m}{y}\theta^{y}(1-\theta)^{m-y} \qquad y = 0,1,2,\ldots\ldots,m$$

Para una muestra aleatoria de tamaño $\boldsymbol{n}$, la función de verosimilitud, está dada por:

$$\ell(y/\theta) = \left(\prod_{i=1}^{n} \binom{m}{y}\right) \theta^{\sum_{i=1}^{n} y_i} (1-\theta)^{mn-\sum_{i=1}^{n} y_i}$$

Donde $y = 0,1, \ldots, m \; \forall i$.

Al aplicar Bayes, la distribución posteriori de $\boldsymbol{\theta}$ dado $\boldsymbol{y}$ es:

$$\boldsymbol{P(y/\theta) \propto \theta^{\alpha+\sum_{i=1}^{n} y_i - 1} (1-\theta)^{\beta+mn-\sum_{i=1}^{n} y_i - 1}}$$

Que tiene forma de

$$B\left(\alpha + \sum_{i=1}^{n} y_i, \beta + nm - \sum_{i=1}^{n} y_i\right)$$

Unas de las ventajas de las distribuciones conjugas es que simplifica enormemente los cálculos

Hay situaciones especiales que tradicionalmente se han usado para simplificar casos y se utilizan como modelos de familias conjugadas. A continuación se menciona el caso de mayor interés para este trabajo como lo es el modelo Beta-Bionomial.

3.9 MODELO BETA-BINOMIAL

Con la distribución binomial-beta se busca modelar el número de éxitos en n ensayos binomial con una probabilidad de éxito p de una Beta (a, b). La distribución Beta por su flexibilidad es considerada como una representación de la aleatoriedad de p, una de las funciones de densidad de una Beta es:

$$f(p|\alpha,\beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} & \text{Para } 0 \leq p \leq 1 \\ 0 & resto \end{cases}$$

Según las posibilidades de éxito conseguiría variar el azar, dependiendo de la aplicación de los ensayos. Por ejemplo, se podría tomar una distribución Beta – Binomial (BB) para modelar:

- El números de motos que se estrellan en una carrera n, el cual el factor predominante no es la destreza de cada conductor, sino la hora en el día.

- El número de titulares de pólizas de seguros de vida que va a morir en un año, donde algunas variables externas (por ejemplo, enfermedad altamente contagiosa, el clima extremo) modera la probabilidad de muerte de toda persona a un cierto grado;

3.9.1. Planteamiento del Modelo Beta-Binomial

Se construye un experimento que se fundamenta en observar $\boldsymbol{n}$ casos independientes, registrando el número de casos favorables que se presentan.

El modelo(o la verosimilitud) se halla, asumiendo que

$$P(Y = y|p) = \binom{n}{y} p^{y}(1-p)^{n-y}, \text{ Para y = 0,1,…., n.}$$

El interés se centra en estimar la proporción $\boldsymbol{p}$ de éxitos dada la muestra. A partir de la perspectiva bayesiana se toma una información previa sobre $\boldsymbol{p}$ que se modeliza mediante la distribución a priori de $\boldsymbol{p}$ se ha supuesto que se recoge por medio de una densidad Beta.

Aplicando el teorema de Bayes,

$$f(p|y) = \frac{f(p)f(y|p)}{f(y)} \propto \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}.\binom{n}{y} p^{y} \propto \binom{n}{y} p^{y}(1-p)^{n-y}$$

Siendo

$$p \in (0,1) \; y \, f\,(p|y) = 0 \;\; para\, p \;\notin (0,1)$$

Dando una función de densidad

$$\boldsymbol{f(y) = f(y\text{:}\, n, \alpha, \beta) = \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \,.\, \frac{\Gamma(y+\beta)\Gamma(n+\beta-y)}{\Gamma(n+\alpha+\beta)} I_{\{0,1,\ldots,n\}}(y)}$$

En efecto, la distribución a posteriori de $\boldsymbol{p}$ sigue una beta de parámetros $(y+\alpha)$ y $(n-y+\beta)$.

SIN EMBARGO, $\boldsymbol{n}$ toma un valor entero no negativo, α> 0, y β> 0, Este modelo se aplica a problemas relativos a proporciones, el cual nos permite tomar experimentos con dos opciones de resultados (uno es de éxito y el otro de fracaso).

La media y la varianza de esta distribución estan dados por:

$$\mu(y) = \frac{\alpha}{\alpha+\beta} \qquad\qquad \sigma_y^2 = \frac{n\alpha\beta(n+\alpha+\beta)}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Se tiene que $\alpha = \beta = 1$, la distribución beta-binomial equivale a una distribución uniforme discreta, como se define en la función de probabilidad.

$$f(y) = \binom{n}{y} = \begin{cases} (n+1)^{-1} & si\ y = 0,1,\ldots,n; \\ \quad 0 & en\ otro\ caso. \end{cases}$$

3.9.1.1. La Media Beta-Binomial

La media está definida como: $\mu(y) = \frac{\alpha}{\alpha+\beta}$

En efecto:

$$\mu(y) = \int_0^1 \mathrm{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \mathrm{y}^{\alpha-1}(1-y)^{\beta-1} dy$$

$$=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}\int_0^1 \mathrm{y}\ \frac{\Gamma(\alpha+\beta+1)}{\Gamma(\alpha+1)\Gamma(\beta)}\mathrm{y}^{\alpha}(1-y)^{\beta-1}d$$

$$=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}.1$$

$$=\frac{\alpha}{\alpha+\beta}$$

3.9.1.2 La Varianza Beta-Binomial

La varianza es definida como:

$$\sigma_y^2=\frac{n\alpha\beta(n+\alpha+\beta)}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$E(y^2)=\int_0^1 y^2 f(y)dy$$

$$=\frac{1}{B(\alpha,\beta)}\int_0^1 y^{(\alpha+2)-1}(1-y)^{\beta}\,dy$$

$$=\frac{\beta(\alpha+2,\beta)}{B(\alpha,\beta)}$$

$$=\frac{\Gamma(\alpha+2).\Gamma(\beta)}{\Gamma(\alpha+2+\beta)}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)}$$

$$=\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

Como

$\sigma^2=var(y)=\alpha_2-\mu^2$ Entonces

$$var(y) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2}$$

$$= \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

Ejemplo 6:

El representante de la marca SONY, anuncia que el 97% de sus electrodomésticos no se tiene que reparar durante el año de garantía. A partirdel punto de vista bayesiano se puede cambiar esta opinión mediante un Be$(4.75, 0.25)$ puesto que:

$$P \sim Be\ (4.75, 0.25)$$

Luego:

Media beta-binomial: $\frac{\alpha}{\alpha+\beta}$

$$\frac{4{,}75}{4{,}75+0{,}25}$$

$$= \frac{4{,}75}{5} = 0{,}95$$

Varianza beta-binomial: $\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$

$$\frac{(4.75)(0.25)}{(4{,}75+0{,}25+1)(4{,}75+0{,}25)^2}$$

$$= \frac{1.1875}{(4{,}75+0{,}25+1)(4{,}75+0{,}25)^2}$$

$$= \frac{1.1875}{(6)(25)} = \ 0{,}00791666$$

Debido a la aparente calidad, se compran 20 electrodomésticos, de los cuales 12 requieren reparación durante el año de garantía. La distribución a posteriori sobre la proporción es entonces ahora

$$\boldsymbol{Be}\,(\mathbf{4.75\ +\ 8, 0.25\ +\ 12})$$

3.10. Estimación Bayesiana**.** Como se explica en algunos textos, los métodos clásicos de estimación se fundamentan solamente en la información que proporciona una muestra aleatoria, mientras que el enfoque bayesiano se fundamenta en la interpretación subjetiva de la probabilidad, el cual considera a ésta como un grado de creencia con respecto a la incertidumbre.

El parámetro es tomado como una variable aleatoria a la que, antes de la evidencia muestral, se establece una distribución a priori, con base en un cierto grado de creencia con relación al comportamiento aleatorio. Cuando se consigue evidenciar la muestra, la distribución *a priori* es modificada y entonces surge una distribución *a posterior.*

Ejemplo 7:

Supóngase que el parámetro $\boldsymbol{p}$ de una binomial tiene distribución a priori de forma:

$$\boldsymbol{\Phi}(\mathbf{p}) = \frac{\boldsymbol{p}^{\boldsymbol{\alpha}-\mathbf{1}}(\mathbf{1}-\boldsymbol{p})^{\boldsymbol{\beta}-\mathbf{1}}}{\boldsymbol{B}(\boldsymbol{\alpha},\boldsymbol{\beta})}$$

Con $\boldsymbol{n}$ ensayos independientes en los que hay $\boldsymbol{x}$ éxito.

Para $\boldsymbol{p}$ y $\boldsymbol{y}$ la distribución conjunta está dada por

$$\boldsymbol{f}(\boldsymbol{p},\boldsymbol{y}) = \binom{\boldsymbol{n}}{\boldsymbol{y}}\boldsymbol{p}^{\boldsymbol{y}}\boldsymbol{q}^{\boldsymbol{n}-\boldsymbol{y}}\frac{\boldsymbol{p}^{\boldsymbol{\alpha}-\mathbf{1}}(\mathbf{1}-\boldsymbol{p})^{\boldsymbol{\beta}-\mathbf{1}}}{\boldsymbol{B}(\boldsymbol{\alpha},\boldsymbol{\beta})}$$

$$= \boldsymbol{f(y|p)\Phi(p)}$$

$$\boldsymbol{f(y) = \int_0^1 \frac{\binom{n}{y} p^{\alpha-1} q^{n-y} (1-p)^{\beta-1}}{B(\alpha,\beta)} . dp}$$

$$\boldsymbol{= \frac{\binom{n}{y}}{B(\alpha,\beta)} \int_0^1 p^{y+\alpha-1}(1-p)^{n-y+\beta-1} . dp}$$

$$\boldsymbol{= \frac{\binom{n}{y}}{B(\alpha,\beta)} B(y+\alpha, n-y+\beta)}$$

De lo anterior, el estimador puntual posterior de $\boldsymbol{\beta}$ será:

$$\boldsymbol{\int_0^1 f(p|y). dp = \int_0^1 \frac{p^{y-\alpha}(1-p)^{n-y+\beta-1}}{B(y+\alpha, n-y+\beta)} . dp}$$

$$\boldsymbol{= \frac{B(y+\alpha+1, n-y+\beta)}{B(y+\alpha, n-y+\beta)} . dp}$$

Es decir:

$$\boldsymbol{\hat{p} = \frac{y+\alpha}{n+\alpha+\beta}}$$

En la estadística clásica (frecuentista) se tendría:

$$\boldsymbol{\hat{p} = \frac{x}{n}} \qquad [1]$$

no tiene en cuenta la distribución a priori.

En este ejemplo se ha puesto que el parámetro $\boldsymbol{p}$ tiene distribución “a priori” beta de parámetros $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$, así como lo indica el ejemplo 6, esta es la mejor hipótesis para $\boldsymbol{p}$.

La estimación se puede realizar para “a priori” de $\boldsymbol{p}$ sin ver los datos es

$$\hat{\mathbf{p}}_{\boldsymbol{o}} = \frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}+\boldsymbol{\beta}} \qquad [2]$$

Que a la vez es la esperanza de una beta. al aplicar estadística frecuentista probablemente se tomaría como estimación [1].

Al aplicar estadística bayesiana varía los elementos; más claro sería

$$\hat{\boldsymbol{p}}_{\boldsymbol{B}} = \left(\frac{\boldsymbol{n}}{\boldsymbol{n}+\boldsymbol{\alpha}+\boldsymbol{\beta}}\right)\left(\frac{\boldsymbol{x}}{\boldsymbol{n}}\right) + \left(\frac{\boldsymbol{\alpha}+\boldsymbol{\beta}}{\boldsymbol{n}+\boldsymbol{\alpha}+\boldsymbol{\beta}}\right)\left(\frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}+\boldsymbol{\beta}}\right) \ [3]$$

Que es la combinación lineal de la estimación muestral y de la estimación “a priori”.

4. ALGUNOS ALGORITMOS DE USO FRECUENTE

Hace algunos años se generaron aplicaciones determinadas de los algoritmos MCMC aplicadas en ciertas áreas de la estadística y también para la reconstrucción de imágenes digitales. Gracias al avance y al desarrollo de nuevas herramientas computacionales apropiadas, se han permitido implantar estos métodos a la inferencia bayesiana, optimización entre otras. Además, han permitido su más práctica solución, ampliando así el alcance de este tipo de métodos (Mackay, 1986).

CÓMO SE ORIGINA:

El algoritmo básico propuesto por (Metropolis et al, 1953) fue generalizado por (Hastings, 1970) dando lugar al conocido algoritmo de Metropolis-Hastings, que constituye la versión más general de la familia de algoritmo MCMC. Más tarde(GemanyGeman, 1984)elaboran un metodo de simulación, que también genera una cadena de Markov y que, tras ser enaltecido en (Gelfand, 1990),pasa a ser conocido en la literatura como muestreador de Gibbs.

Para profundizar (Congreso Nacional de Estadística e Investigación Operativa, 2003), o también en el link http://web.udl.es/usuaris/esi2009/treballs/15_1_1.pdf)

4.1. EL ALGORITMO DE METRÓPOLIS-HASTINGS

(Metropolis, 1953) relaciona la simulación de un parámetro del que se conocen la densidad a posterior $f(\theta|y)$ posiblemente, la constante de integración. Este algoritmo es semejante al algoritmo del rechazo para simular variables aleatorias. Con este proceso, se busca una densidad condicional $g(.|\theta)$ o la proporción de distribución que sea fácil muestrear. Posteriormente se crean observaciones de esta distribución de propuesta para decidir si pertenecen a la distribución de $\theta|x$ mediante un sorteo.

Comprobación de la convergencia:

I. Comparación de los resultados con valores iníciales diferentes.
II. Gráficos de los valores θ (t).
III. Gráficos de la media estimada (running means) de θ.
IV. Diagnósticos formales de convergencia.

Algoritmo: Metropolis – Hastings

1. Se inicializa con un valor $\theta^{(0)}$.
2. Se considera una densidad $q\left(\cdot\,|\theta^{(j-1)}\right)$de la cual se genera $\phi_{t.}$
3. Se evalúa la probabilidad:

$$\alpha(\theta^{(j-1)}, \phi_t)= \min\left\{1, \frac{\pi(\phi_t)q(\phi_t|\theta^{(j-1)})}{\pi\left(\theta^{(j-1)}\right)q(\theta^{(j-1)}|\phi_{t.})}\right\}$$

4. Se acepta ϕ_t con probabilidad α. Si ϕ_t es aceptada entonces $\theta^{(j)} = \phi_t$ de otra forma $\theta^{(j)} = \theta^{(j-1)}$.
5. Continuar con 2.

4.2. ALGORITMO MCMC (MARKOV CHAIN MONTE CARLO)

- “. . .Muestrear la distribución a posteriori y calcular la cantidad a posterior de interés mediante MCMC son los retos más importantes de la computación bayesiana más avanzada .” (Chen Shao e Ibrahim, 2000)

Debido a la revolución de las Cadenas de markov de Monte Carlo (MCMC) en los 90 permitió que la inferencia bayesiana redujera el uso de distribuciones conjugadas a las verosimilitudes de: Ratios y proporciones (Bernoulli/binomial); conteos (Poisson), medias, varianzas y regresiones de variables cuantitativas (Normal y derivadas)

Las cadenas MCMC (Markov Monte Carlo) son métodos o clases de algoritmos para simular datos autocorrelacionados que se extraen de las distribuciones de probabilidad. Esta cadena maneja estimaciones empíricas para crear inferencias en distribuciones multidimensionales que se presentan.

En la estadística bayesiana; uno de los procedimientos de MCMC es solucionar y generar muestras para parámetros de interés, que al realizar simulación de MCMC con distribucion estacionaria dada por la distribuciòn a posterior $f(\theta|y)$(Gill, 2002).

Además , MCMC no es utilizado simplemente para inferencia bayesiana, sino también para casos en los que, por ejemplo, la verosimilitud no puede ser calculada claramente debido a información faltante o dependencia compleja (Geyer, 2011).

Se requiere conocer la distribución $f(\theta|y)$, donde θ ∈ Θ es el vector de parámetros, y $\mathbf{\Upsilon}$ son los datos. La idea de trabajar con cadenas de Markov es simular un proceso de Markov en Θ, el cual converja a la distribución estacionaria $f(\theta|y)$.

En la Estadística Bayesiana el problema es el inverso: cómo construir la cadena de Markov para obtener una muestra de una distribución dada. Existen dos algoritmos que son utilizados como complemento a este método ellos son:

El Gibbs Sampler y los Metropolis - Hastings, están relacionados en el paquete especializado WinBUGS (Bayesian Analysis Using the Gibbs Sampler) en entorno

Windows según (Kery, 2012). Este programa ha sido desarrollado para problemas de inferencia estadística bayesiana.

Las MCMC se basan en el siguiente algoritmo:

Algoritmo[1] (Aumentación de datos)

Dado $\boldsymbol{y^n} = (\boldsymbol{y_1^n}, \boldsymbol{y_2^n}\,)$

Repetir

Generar

$$\boldsymbol{y_1^{n+1} \sim g_1(y_1|y_2^n)}$$

$$\boldsymbol{y_2^{n+1} \sim g_2(y_2|y_1^n)}$$

Imcrementar $\boldsymbol{n}$

Algoritmo MCMC:

$$\theta^{(0)} \leftarrow x$$

$$for\ i = 1\text{to M}$$

$$\theta^{(i)} = f\left(\theta^{(i-1)}\right)$$

Fin

PROCESO

El muestreador Gibbs es una variante del algoritmo de Monte Carlo mediante Cadenas de Markov, en el cual se optan por tomar muestras al azar de una distribución de probabilidad que cumple con la Condición de Markov para aproximar, mediante la toma de muestras aleatorias sobre un modelo inicial, la distribución de probabilidad conjunta de la variable observada (Gagnon, 2011). Así, el algoritmo muestrea un nuevo valor de la variable, de acuerdo a una distribución de probabilidad basado en los valores previos.

Este proceso se repite de forma iterativa y los nuevos valores actualizan la distribución de probabilidad con cada iteración para continuar el muestreo. Al terminar el proceso iterativo, la secuencia de Gibbs debería converger a una distribución estacionaria o de equilibrio, independiente de los valores iniciales. Esta distribución estacionaria, por construcción, debería ser la distribución objetivo que se está tratando de simular (Tierney, 1994). Véase (Revista de la Facultad de Ingeniería U.C.V, 2014)

4.3 MUESTREADOR DE GIBBS.

El algoritmo gibbs sampling, (tsay, 1994) estima la distribución a posteriori de estos parámetros que se aplican los métodos mcmc, los cuales crean una cadena de markov irreducible y aperiódica para la cual la distribución permanente es igual a la distribución a posteriori de interés(albert, 2007).

El muestreador de Gibbs (Geman&Geman);(Gelfand & Smith, 1990);(George, 1992); (Stephens y Smith, 1992)es un método iterativo que se ha implementado en la toma de muestras de una distribución posterior $\boldsymbol{p}(\boldsymbol{\theta}|\boldsymbol{y})$ para un vector de parametros desconocidos. Ademas, El muestreador de Gibbs puede verse como un caso particular del algoritmo de Metropolis-Hastings, sin embargo, como destacan(Robert y Casella, 1999) tiene algunas caracteristicas particulares que le dan entidad propia.

Sea $\theta=(\theta_1,\dots,\theta_k)'$ el vector paramétrico y sean $p(\theta_1|\theta_2,\dots\dots,\theta_k)$, $p(\theta_2|\theta_1,\dots\dots,\theta_k)\dots.p(\theta_k|\theta_2,\dots\dots,\theta_{k-1})$las distribuciones condicionales completas obtenidas a partir de la distribución final (θ).

El algoritmo de Gibbs parte de un punto inicial $\theta^{(0)}=(\theta_1^{(0)},\dots,\theta_k^{(0)})$ que comienzan la cadena $\theta_1^{(1)}$

A partir de $p\left(\theta_1\middle|\theta_2^{(0)},\dots,\theta_k^{(0)}\right)$, $\theta_2^{(1)}$

a partir de $p(\theta_1^{(0)},\dots,\theta_k^{(0)})$

y así como $\theta_k^{(1)}$a partir de $p\left(\theta_k\middle|\theta_1^{(1)},\dots,\theta_{k-1}^{(1)}\right)$[1]

la relación estocástica entre los muestreos subsecuentes está descrita por las probabilidades de transición markovianas de primer orden:

$\pi\left(\theta^{(m)},\theta^{(m+1)}\right) = \prod_{j=1}^{k} p\left(\theta_j^{(m+1)}\middle|\theta_m^{(h)}\ para\ h > j, \theta_h^{(m+1)}\ para\ h < j, y\right)$[2]

El cual es el producto de probabilidades condicionales formado a partir de $(\theta_k|\theta_1^1, \ldots, \theta_{k-1}^1)$. De manera general, la sucesión $\{\theta^{(i)}\}$, así creada es una realización de una cadena de Markov con la distribución final (Gelfand y Smith, 1990). Para profundizar (Estadística Española Vol. 35. Núm. 134, 1993, págs. 629 a 644)

4.3.1. Ventaja del Muestreador de Gibbs: No obstante si la dificultad sea multidimensional, las distribuciones de las cuales es necesario generar valores unidimensionales,hallando una variedad de métodos de simulación exacta para distribuciones univariantes (véase ejemplo (Ripley, 1987).

De manera general el algoritmo procede:

1. Inicializar el conteo de las iteraciones de la cadena con $j = 1$ y el conjunto de valores iniciales $\theta^{(0)}$= $(\theta_1^{(0)}, \ldots, \theta_k^{(0)})'$.

2. Obtener un nuevo valor $\theta^{(j)}$= $(\theta_1^{(0)}, \ldots, \theta_d^{(0)})$ a partir de $\theta^{(j-1)}$a través de una generación sucesiva de valores

$$\theta_1^{(j)} \sim p\left(\theta_1\middle|\theta_2^{(j-1)} \ldots, \theta_d^{(j-1)}\right)$$

$$\theta_2^{(j)} \sim p\left(\theta_2\middle|\theta_1^{(j)}, \theta_3^{(j-1)} \ldots, \theta_d^{(j-1)}\right)$$

...

$$\theta_d^{(j)} \sim p\left(\theta_d\middle|\theta_1^{(j)}, \ldots, \theta_{d-1}^{(j)}\right)$$

3. Cambiar el contador de j a $j + 1$ y volver al paso 2 hasta que la cadena converja. Cuando la cadena converja, el valor resultante θ^j es un muestreo

de p. Conforme el número de iteraciones se incremente, la cadena se aproximar a su condición de equilibrio.

ALGORITMO 2.

- Paso 0. Valores iníciales $\theta^{(0)} = (\mu_0, h_0)$

- Paso 1. Para obtener $\theta^{(1)} = (\mu_1, h_1)$:

  Se muestrea $\mu_1$de $\pi(\mu|h = h_0, y)$,

  (Se genera un valor de la dist. Normal)

  Se muestrea h_1 de $\pi(h|\mu = \mu_1, y)$

  (Se genera un valor de la distancia Gamma)

  Se actualiza (μ_0, h_0) a (μ_1, h_1),

- Paso K. Actualizar $\theta^{(k)}$=(μ_k, h_k), a partir de $\theta^{(k-1)}$

Después de N realizaciones: $\theta^{(0)}, \theta^{(1)}, \dots, \theta^{(N)}$,Se obtiene que $\{\theta^{(t)}\}$ es una cadena de Markov cuyas propiedades de transición son

$p\left(\theta^{(t+1)}\middle|\theta^{(t)}\right) = \prod \pi\left(\theta_{t+1i}\middle|\theta_{tj}, j > i, y\right)$,de donde, $\theta^{(t)} \to \theta \sim \pi(\theta|y)(t \to \infty)$

Ver(Robert Ch., 1996)

Así, para N suficientemente grande…

Luego:

←La serie $\theta^{(0)}, \theta^{(1)}, \ldots, \theta^{(N)}$ puede analizarse casi como una muestra independiente de $\pi(\theta|y)$, y por tanto, cantidades muéstrales estimarán las cantidades a posteriori respectivas (media muestral para la media a posteriori, cualquier momento o percentil muestral para el correspondiente a posteriori, o la curva descrita por el histograma de valores para un parámetro θ_i aproxima la forma de la curva de la distribución marginal $\pi(\theta_i|y)$).

5. SOFTWARE

5.1. WINBUGS – OPENBUGS

WinBUGS es un software fundamentado en lenguaje de programación que se utiliza para generar una muestra al azar de la distribución posterior de los parámetros de un modelo Bayesiano, y igualmente permite simular la distribución a posterior combinando varias técnicas MCMC. El usuario simplemente tiene que definir la estructura del modelo en cuestión. Además, WinBUGS se pueden ejecutar incluso desde otros paquetes de software como por ejemplo R, Matlab y Excel.

Open BUGS es un software informático para el análisis bayesiano de complejos modelos estadísticos utilizando la cadena de Markov Monte Carlo (MCMC).

Con este programa se pueden realizar trabajos de métodos bayesianos aunque han surgido ciertos inconvenientes como lo es la carencia de programas o software especializados; pero, existen paquetes estadísticos como: SAS, SPSS, entre otros, que tienen módulos para realizar estadística bayesiana. Además, el internet ofrece programas de acceso gratuito que consiente en realizar simulaciones basadas en cadenas de Markov de forma simple y efectiva con variedades de modelos llamados BUGS, que es un acrónimo de Bayesian análysis Using the Gibbs Sampler. Este programa está disponible en www.mrc-bsu.ac.uk/bugs , fue desarrollado UK Medical Research Council y el Imperial College of Science, Technology and Medicine. Sin embargo, hay otros programas que también desarrollan problemas bayesianos como: FirsBayes, el BACC, y el R que dentro de este paquete existe soluciones a ciertos problemas, por ejemplo el MCMCPack y CODA.

Se puede tener una idea sobre este programa donde el lenguaje de programación es utilizado para generar muestras de una distribución posterior de los parámetros de un modelo. Con este programa se generan ciertas operaciones como lo son: Operaciones binarias con comandos File, Windows, Help; Operaciones de edición con comandos Tools, Edit, Text.; Funciones MCMC.

Con comandos Info, Model, Inference, Optims ; Operaciones especializadas con comandos Dooble, Maps.

Ejemplo:

```
MODEL {
  p ~ dbeta(a,b)
  x ~ dbin(p,n)
}
DATA list(a=1,b=1,x=11,n=25)
```

5.2. CARACTERÍSTICAS DEL SOFTWARE

Según Koop (1999) numera ciertos requisitos básicos que todo software bayesiano será:

- Deberá ser computacionalmente eficiente
- El grupo de soporte debe ser amplio y reconocido.
- Debe proporcionar simuladores posteriores para la clase de modelos que los investigadores quieran usar.
- Debe permitir al interesado graficar la a posterior y la a priori.
- Debe permitir al usuario realizar un análisis de sensibilidad a priori de una manera fácil.
- Entre otros.

Lo que se requiere es centrarse en el cálculo de las distribuciones posteriores $f(\theta|y)$del parámetro θ dado los datos observados$\boldsymbol{y}$**,** como se observa:

$$f(\theta|y) = \frac{f(\theta)f(y|\theta)}{f(y)} \propto f(y|\theta)f(\theta)$$

En la formula, la distribución posterior contempla la distribución a priori como también la información de los datos (en la $f(\theta)$) y en la verosimilitud.

$$f(y|\theta)=\prod_{i=1}^{n} f(y_i|\theta)$$

6. CONCLUSIONES

Se puede notar, que la estadística bayesiana permite realizar inferencias combinando la información a priori sobre el suceso que trabajamos y la información muestral. La combinación de información puede efectuarse a partir de datos puntuales o considerando la información como una función de distribución. Esta última forma facilita la toma de decisiones basadas en un modelo final que integra las fuentes de información disponibles, dando como conclusión una inferencia más completa. La investigación desarrollada con este tipo de estadística exige una mayor contribución entre los datos y la ayuda tecnológica. El investigador debe aportar toda la información a priori disponible sobre el parámetro acerca del cual desea hacer inferencias, y esta información debe ser modelable matemáticamente.

Gracias al avance y al desarrollo de nuevas herramientas computacionales apropiadas, se han permitido constituir ciertos métodos como práctica solución y ampliando así el alcance de los algoritmos MCMC como aplicación en ciertas áreas de la estadística, como también para la reconstrucción de imágenes digitales.

Para finalizar diremos que la estadística frecuentista y bayesiana:

- ✓ Al establecerse una comparación entre estas dos vertientes se debe tener clara sus diferencias y ventajas.
- ✓ No es posible realizar inferencias frecuentistas e interpretarlas de forma bayesiana
- ✓ Usar la conjugada apropiada sino es posible este paso, usar un algoritmo como Metrópolis o preferiblemente MCMC (la cadena de Markov Monte Carlo)

REFERENCIAS

Albert, J. (2007). Bayesian Computatin with R. Revista de Métodos Cuantitativos para la Economía y la Empresa, 50 – 70.

Agresti Alan; Brent A. Coull, Approximate Is Better than "Exact" for Interval Estimation of Binomial Proportions. The American Statistician, Vol. 52, No. 2. (May, 1998) 119-126.

Agresti Alan and Yongyi Min, On Small-Sample Confidence Intervals for Parameters in Discrete Distributions. Biometrics, Vol. 57, No. 3 (Sep., 2001)
963-971

Agresti, A. &Caffo, B. (2000), 'Simple and Effective Confidence Intervals forProportion and Differences of Proportions Result from Adding Two Successesand Two Failures', The American Statistician 54, 280–288.

Agresti, A. &Coull, B. (1998), 'Approximate is Better than Exact for Interval
Estimation of Binomial Proportions', The American Statistician 52, 119–126.

Agresti, A. & Min, Y. (2001), 'On Small-Sample Confidence Intervals for Parameters in Discrete Distribution', Biometrics 57, 963–971.Canavos George, Probabilidad y Estadística. McGraw-Hill, México, 1988.

Bayes, T. (1763). «An Essay towards solving a Problem in the Doctrine of Chances.». Philosophical Transactions of the Royal Society of London 53: 370–418. doi:10.1098/rstl.1763.0053.

Bernardo JM. Intrinsic credible regions.An objetcive Bayesian approach to interval estimation.Test 2005; 14(2): 317-384 (disponible en http://www.uv.es/~bernardo/2005Test.pdf).

Bernardo, J. M. and A. F. M. Smith (1994). Bayesian Theor. Chichester: Wiley.

Bolstad, William M. (2004). Introduction to Bayesian Statistics, John Wiley ISBN 0-471-27020-2

Box, G. E. P. and G. C. Tiao (1973). Bayesian Inference in Statistical Analysis. Reading, Massachusetts: Addison-Wesley

Brown, L., Cai, D. &DasGupta, A. (2002), 'Confidence Intervals for a Binomial Proportion and Asymptotic Expansions', The Annals of Statistics 30, 160– 201

Canavos, G. (1988), Probabilidad y Estadística: Aplicaciones y métodos, McGraw Hill, México D. F., México.

Casella, G. and R.L Berger (1990). Statistical Inference. Belmont, California: Duxbury Press

Congreso Nacional de Estadística e Investigación Operativa. Lleida, del 8 al 11 de abril de 2003

Correa, J. C. (2005). Estadística Bayesiana. Medellín: Universidad Nacional - Sede Medellín.

Clavijo Jairo Alfonso, Estimación de Proporciones Multinomiales. Universidad del Tolima, 2005.

Geman. S. & Geman, D. (1984),"Stochastic relaxation, Gibbs distribution and the Bayesian restretion of imagines", IEEE Trans. On Pattern Analysis and Machine Inteligence 6, 721 - 741

Gelfand, A.E. y Smith, A.F.M. (1990). “Sampling based approaches to calculating marginal densities”. Journal of the American Statistical Associatin, 85, 398-409.

Gilks, W. R.;Richardson, S.; Spiegelhalter, D. J. Markov Chain Monte Carlo in Practice. Londres: Chapman and Hall, 1996. [Una introducció excel. lent als mètodes MCMC i llursaplicacions]..

Goodman Leo A.,On Simultaneous Confidence Intervals for Multinomial Proportions. RevistaTechnometrics Vol. 7, No. 2 (1965) 247-254.

Hastings, W.K. (1970).“Monte-Carlo sampling methods using Markov chains and their applications”.Biometrika, 57, 97—1

Jack C. Lee and Y.L.Lio .”A note on Bayesian Estimation and Prediction for the Beta – Binomial model.

Jeffreys, H. (1961).Theory of Probability, 3rd ed. Oxford Classic Texts in the Physical Sciences. Oxford Univ. Press, Oxford. MR1647885

K´ery, Marc and Michael Schaub (2012). Bayesian Population Analysis Using WinBUGS: A Hierarchical Perspective. Academic Press. 556 pp

Lawrence D. Brown, T. Tony Cai and Anirban DasGupta, *Interval Estimation for a Binomial Proportion*. Revista Statistical Science Vol. 16, No. 2 (2001) 101-133.

Lehmann, E. L. and G. Casella (1998). *Theory of Point Estimation* (Second ed.). London: Springer

López de Castilla, C. (2011). “Estadística Bayesiana” Universidad Nacional Mayor de San Marcos. EP4066. 2011-2. August 18, 2011

Mackay, D.J.C.: Introduction to Monte Carlo Methods. Springer (1986)

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. y Tefler, E. (1953)

McCulloch, R.E. y Tsay R.S. (1994a).“Bayesian analysis of autoregressive time series via the Gibbs sampler”.Journal of Time Series Anal ysis, 15, 235—250
Revista Colombiana de Estadística Diciembre 2008, volumen 31, no. 2, pp. 211 a 228

Quesen berry C.P. and D.C. Hurst,*Large Sample Simultaneous Confidence Intervals for Multinomial Proportions*. Technometrics Vol. 6, No. 2 (1964) 191-195.

ANEXOS

Anexo A. Probabilidad a priori: beta

Probabilidad a priori: beta

La función de probabilidad de la proporción diseñada adquiere una distribución beta, cuya expresión es:

$$\pi = Be\ (\alpha, \beta)$$
$$f(\pi) = \lambda \pi^{\alpha-1}.(1-\pi)^{\alpha-1}$$

Siendo λ la constante de proporcionalidad que hace que f sea densidad. Para asignar los parámetros α, β que definen la función, es habitual utilizar la expresión de los siguientes estadísticos:

– La media de cualquier distribución beta es igual a $E(\pi) = \frac{\alpha}{\alpha+\beta}$;

– La variancia de una distribución beta es igual a: $\sigma^2 = \frac{\alpha.\beta}{(\alpha+\beta)^2.(\alpha+\beta+1)}$

Estas dos ecuaciones no lineales nos permiten obtener los valores concretos de $\alpha\ y\ \beta$.

Anexo B. Probabilidad posterior $f\left({}^{\Pi}/_{x}\right)$

Probabilidad posterior $\boldsymbol{f\left({}^{\Pi}/_{x}\right)}$
De acuerdo con la fórmula de Bayes para el caso continuo:
Inf. *A posteriori* α datos muestrales × inf. *a priori*
$$f\left({}^{\Pi}/_{x}\right) \propto f\left({}^{x}/_{\pi}\right) f(\pi)$$
$$f\left({}^{\Pi}/_{x}\right) = \binom{n}{x} \pi^{x}.(1-\pi)^{n-x}.\lambda\pi^{\alpha-1}.(1-\pi)^{\beta-1}$$
$$f\left({}^{\Pi}/_{x}\right) \propto \pi^{x+\alpha-1}.(1-\pi)^{n-x+\beta-1}; 0 < \pi < 1$$
Que alcanza una distribución Beta de parámetros $\alpha + x,\ \beta + n - x$.

Aplicación de Winbugs

Este programa parece extraño cuando se usa por primera vez. Resulta poco evidente ya que no procede de forma lineal, sino que requiere de muchos pasos que parecieren ser repetitivos pero que no lo son.

Dentro de este programa ciertos símbolos tienen un significado especial. Por ejemplo: todos los nombres de distribución y probabilidad comienzan con la letra "d" (por "distribución"), además el símbolo ~ significa "se distribuye como" y es aplicado para especificar la distribución de los datos, es decir, para especificar la distribución a priori.

Para esta situación tendríamos en cuenta, la distribución binomial con parámetros $\boldsymbol{n}, \boldsymbol{p}$ que se define $\boldsymbol{dbin}(\boldsymbol{p}, \boldsymbol{n})$, con proceso de Bernoulli con parámetro $\boldsymbol{p}$.

Por ejemplo: El número de lanzamientos de 100 monedas para obtener caras, y tiene una distribución $\boldsymbol{dbin(0.5, 100)}$**;** el número de bolas negras de una caja, con una muestra de tamaño n, en la que la proporción de bolas negras es $\boldsymbol{p}$, y tiene una distribución $\boldsymbol{dbin(p, n)}$.

Para Realizar Inferencia Sobre Proporción.

dbeta (a, b) es la distribución beta con parámetros ay b. La distribución dbeta es de una familia muy flexible; se aplica a una cantidad desconocida que toma valores entre 0 y 1.Por ejemplo, Un importante caso especial es dbeta (1, 1), es la distribución uniforme (plana) previa sobre el intervalo (0,1). Sin embargo, la distribución dbeta (0, 0) es una distribución inadecuada con zona de curva infinita, y en la práctica se utiliza dbeta (ε, ε) con ε un pequeño número como 0.001.

Realizaremos el siguiente Ejercicio

Un profesor en clase de estadísticas lanzó una moneda n = 25 veces y la "cara" se observada x = 11 veces. Para obtener la distribución posterior de p, Se podría establecer una cantidad desconocida de p (la cantidad de veces que la moneda cae cara), o se podría tener algún conocimiento previo. En cualquier caso, es más conveniente tener en cuenta lo anterior con una distribución beta.

Ejemplo:

Para una $\mathrm{p} \sim \mathrm{dbeta}\ (1, 1)$. Con WinBUGS u OpenBUGS se necesita establecer la probabilidad de los datos de x. Donde x es el número de caras, con una muestra de tamaño $n = 25$, la probabilidad de x éxitos con n observaciones de una distribución Bernoulli, se específica de esta manera:

$$\boldsymbol{x \sim dbin\ (p, n)}$$

Generar el Modelo:

```
MODEL {
  p ~ dbeta(a,b)
  x ~ dbin(p,n)
}
DATA list(a=1,b=1,x=11,n=25)
```

Elaborar el Modelo y Realizar Simulación de Valores

Posteriormente al escribir el código del modelo completo, los datos y los valores iniciales en un archivo tenemos que compilar y ejecutar el modelo.

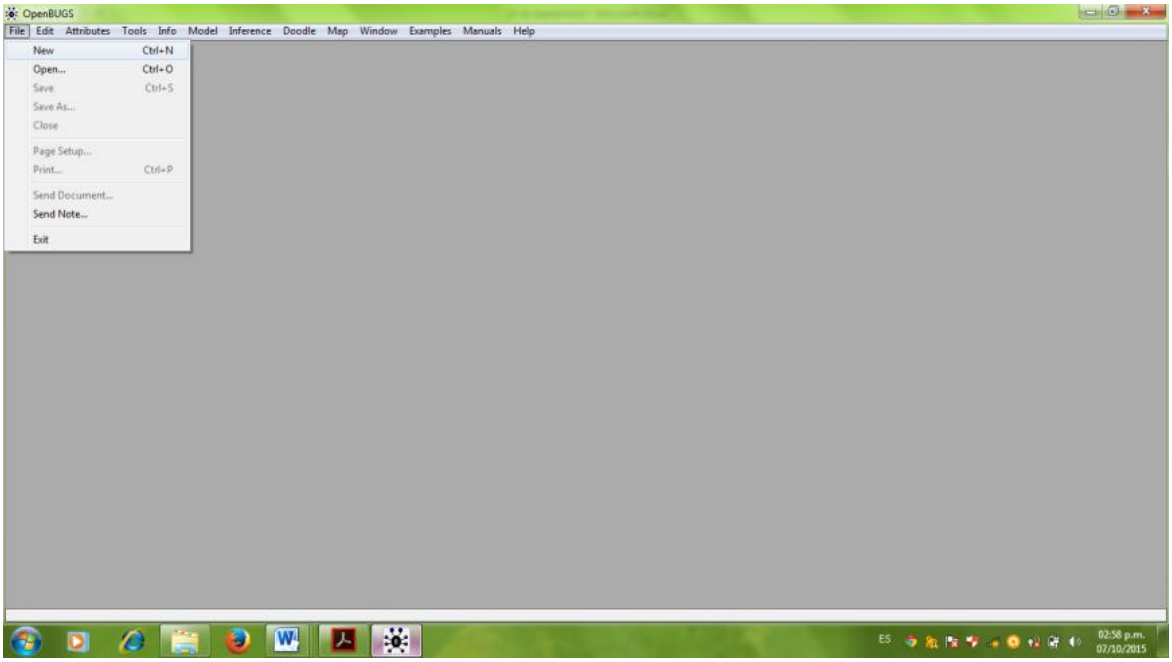
Este procedimiento se resume mediante los siguientes pasos:

1. Escribir modelo
2. Establezca las condiciones iniciales (opcional)
3. No se olvide de guardar la secuencia de comandos antes de ejecutar!
4. Abrir **Model Specification Tool** (Model > Specification)
5. Elija “model”, clic “Check Model”
6. Resalte los datos, clic “Load Data”
7. Conjunto de cadenas (typically 3-5)
8. Clic “compile”

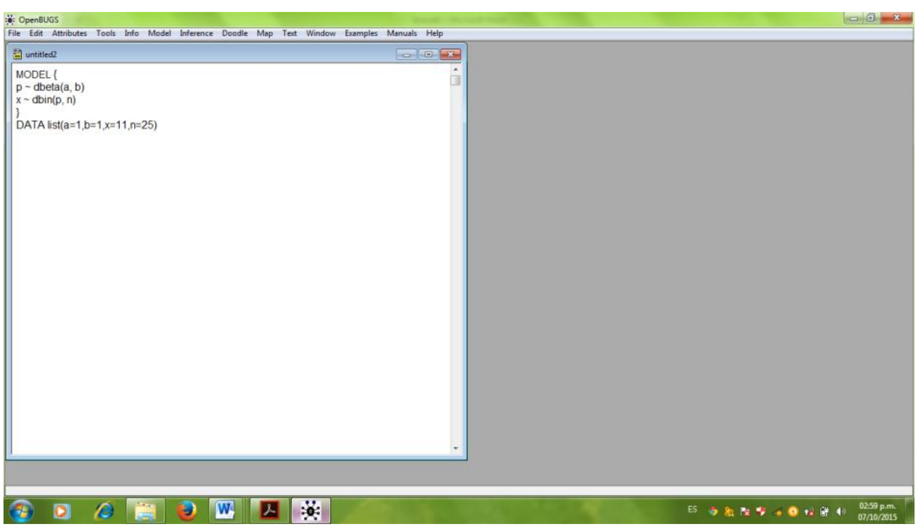
9. Luego, Clic “Gen inits” y / o resaltar las condiciones iníciales y haga clic “load inits”
10. Abra **Sample Monitor Tool** (Inference > Samples)
11. Especificar variables para rastrear (introducir nombre de la variable en la ventana, haga, clic “set”)
12. Abrir **Update Tool** (Model > Update)
13. Clic “update” para ejecutar sampler
14. Use **Sample Monitor Tool** para evaluar el ajuste del modelo, use **UpdateTool** para funcionar más tiempo si es necesario.

Pasos a Seguir Para la Ejecución de Winbugs

- Abrimos WinBUGS, File>New



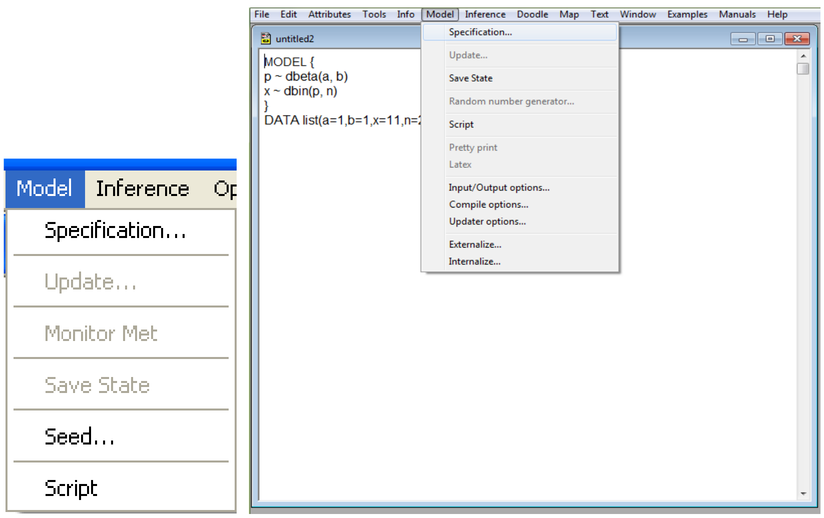
Y escribimos (o pegamos) el modelo, los datos y los valores iniciales en la ventana que aparece.



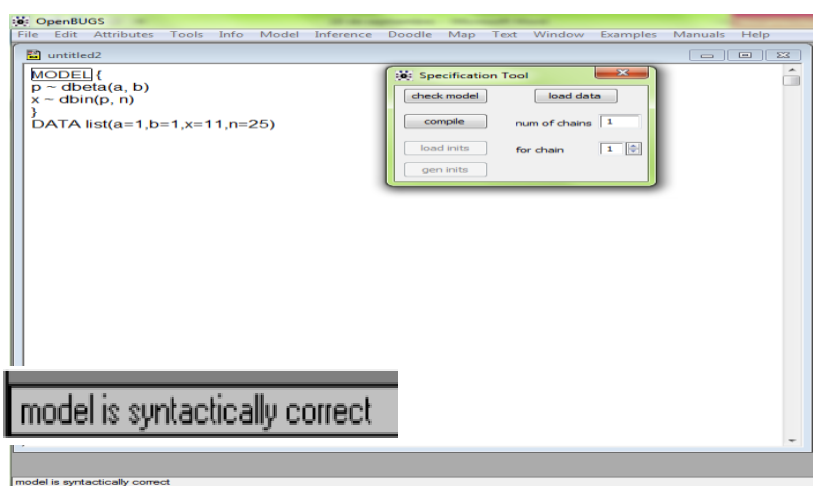
Luego seleccionamos Model>Specification

Esta herramienta contiene todas las operaciones básicas necesarias para inicializar el algoritmo MCMC (comprobación de los modelos " sintaxis del código, cargar los datos, compilar el modelo y establecer los valores iniciales) y especifique el número de cadenas que queremos generar.

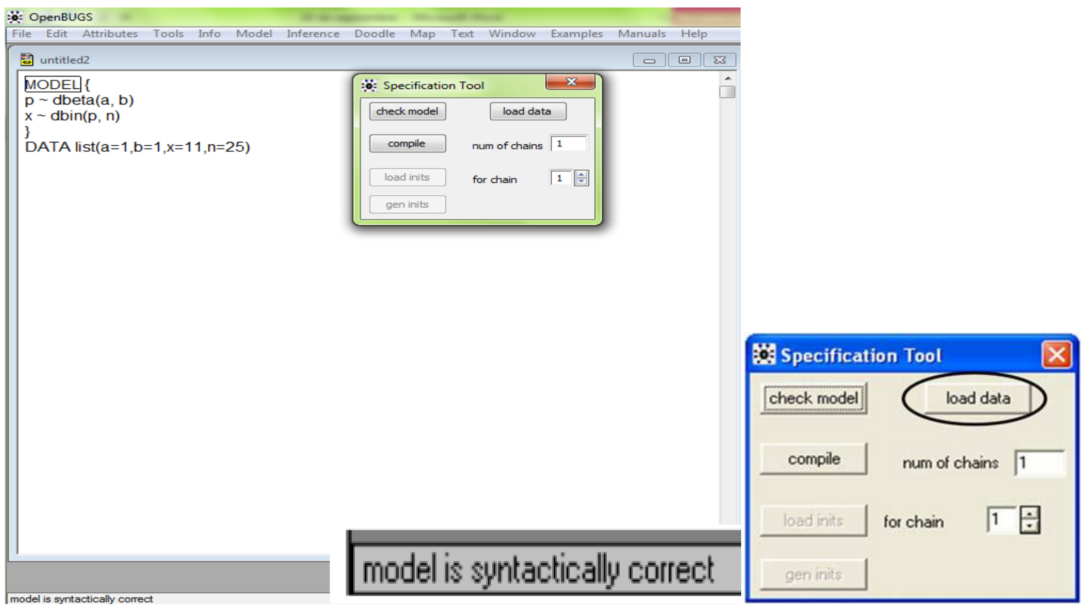
- Luego, con el puntero del mouse seleccionamos el texto del modelo y seleccionamos chek model en la ventana Specification Tool.



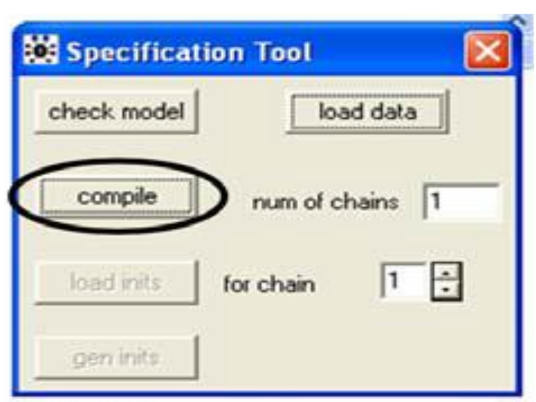
Si el modelo está bien escrito, aparecerá el mensaje model is syntactically correct en la esquina inferior izquierda.



- ***load data* (Cargar los datos):** seleccionamos con el puntero del mouse los datos y hacemos clic en **load data** de la ventana Specification Tool. Este comando ***load data*** permite cargar los datos y se activa una vez que un modelo ha sido comprobado con éxito, y deja de ser activa una vez que el modelo ha sido compilado correctamente.

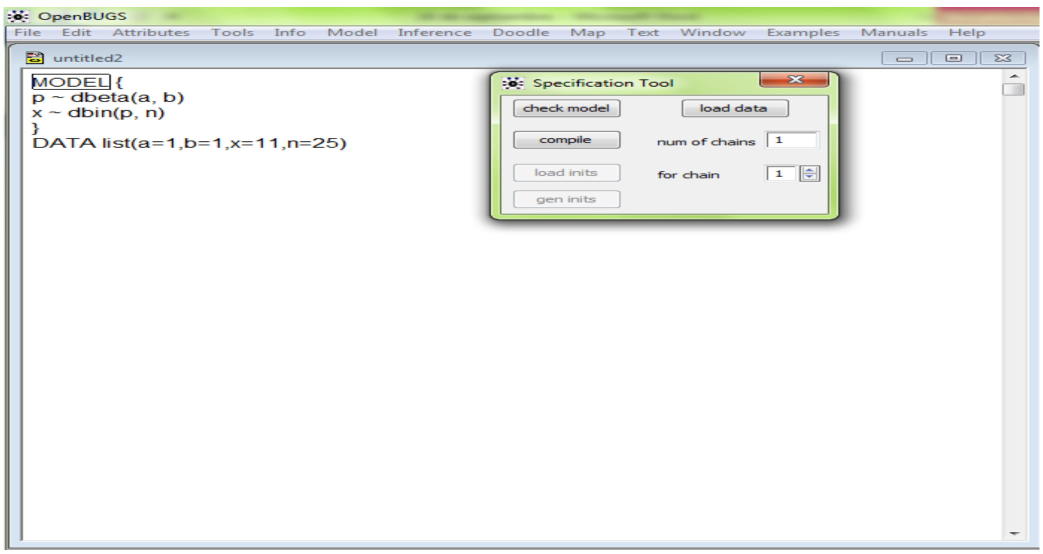


- Luego se cargan todos los datos, después, hacer clic en "Compile" en la ventana de Specification Tool

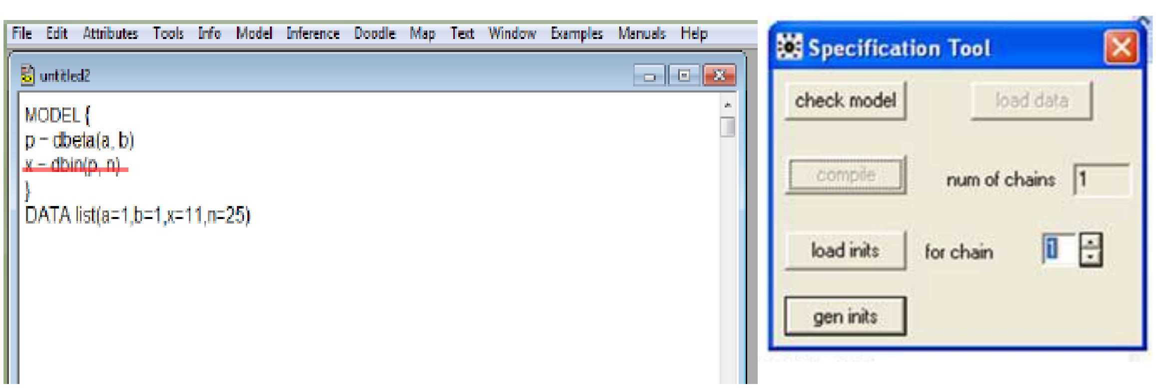


Compile: Este comando “compile” permite la compilación y construye las estructuras de datos necesarios para llevar a cabo el muestreo de Gibbs. El modelo es verificado su integridad y coherencia con los datos.

- Si la compilación tiene éxito, entonces el "modelo compilado" aparecerá un mensaje en la barra inferior aparecerá un mensaje de model compiled

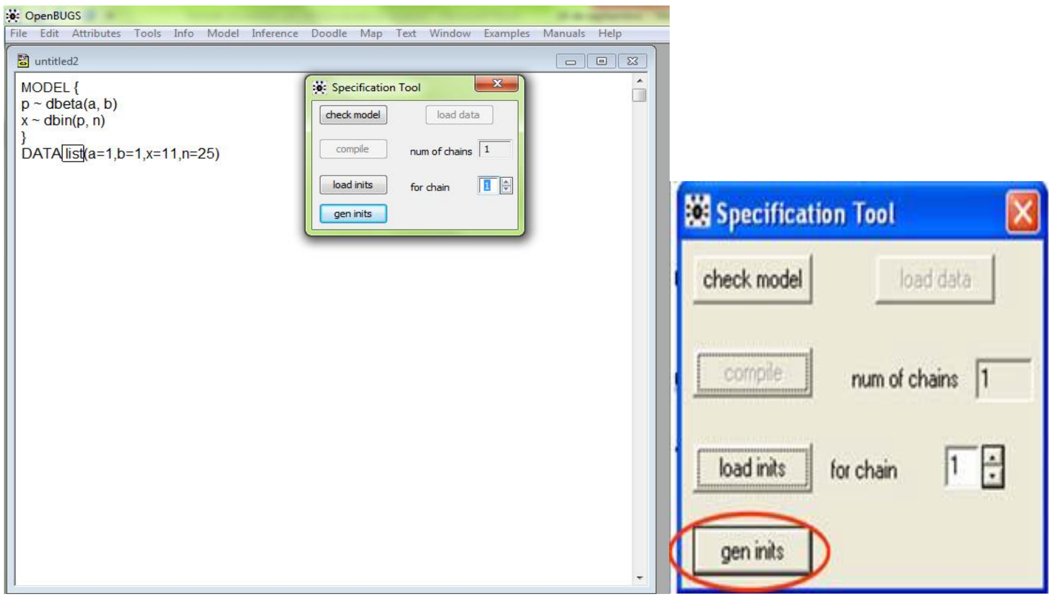


- seleccionamos con el puntero del mouse los valores iníciales y hacemos clic en load inits;

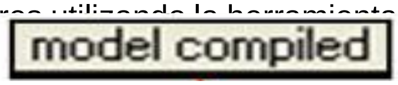


El comando load inits permite cargar los valores iniciales para la cadena indicada sin embargo se puede ejecutar una vez que se ha iniciado el muestreo Gibbs. Se tendrá el efecto de iniciar la toma de muestras en una nueva trayectoria.

Luego, hacemos clic en gen ints

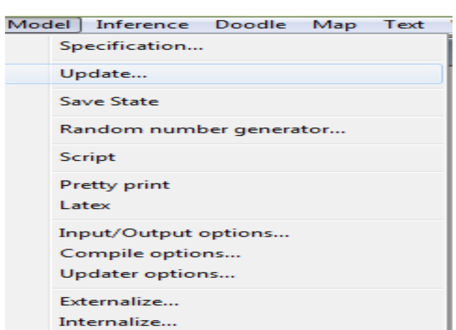


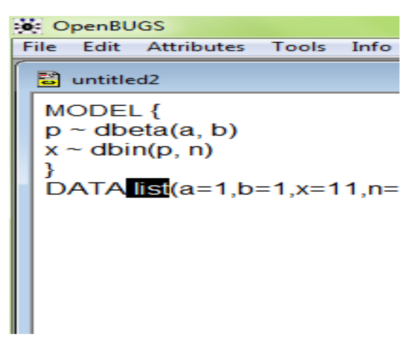
"genints" es útil cuando el número de nodos estocásticos es grande; por ejemplo en los casos de los valores perdidos, efectos aleatorios o datos latentes. Después de ajustar (carga o generación) los valores iniciales el algoritmo se inicializa y está listo para simular valores utilizando la herramienta de "actualización", que se describe en el siguiente paso



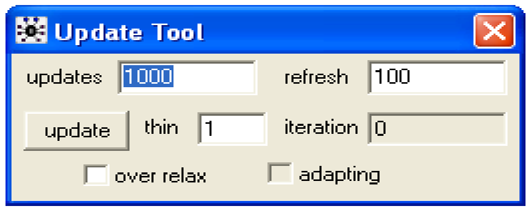
En este momento **Supervisaremos los nodos**

Ahora vamos a **Model>Update** y aparece la ventana **UpdateTool**.

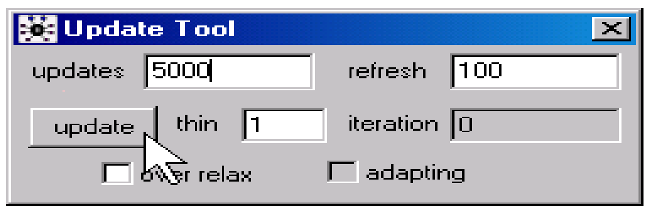




Cuando tenemos un gran número de iteraciones (simulaciones), volvemos a la ventana Sample Monitor Tool. Por ejemplo, 1000 iteraciones.



O como también pueden ser 5000 iteraciones



El cuadro anterior tiene las siguientes características:

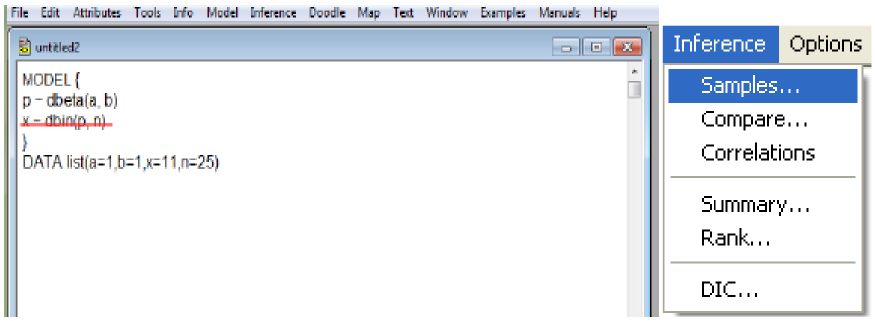
a. **Updates**: Se refiere al número de iteraciones adicionales queremos generar.

b. **Refresh** Se refiere al número de iteraciones thatWinBUGS utilizarán para actualizar el índice de iteración actual en la herramienta de "update" y actualizar las parcelas de trazas en línea que están disponibles a través de la función "Sample monitor" tool.

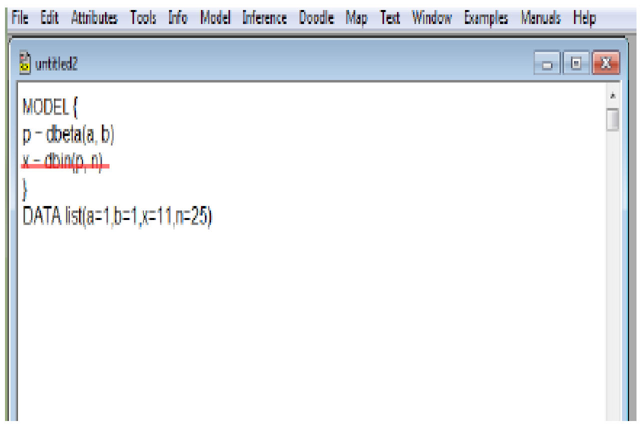
c. **Update**: Al pulsar este botón va a generar iteraciones adicionales igual al número definido en el cuadro de "Updates".

d. **Thin**: Define la delgada (lag) de iteraciones mantienen. El WinBUGS generará $k \times T$ iteraciones pero almacenará sólo el último de cada secuencia de valores $k$generado.

- Seleccionamos con el puntero del mouse Inference>Sample

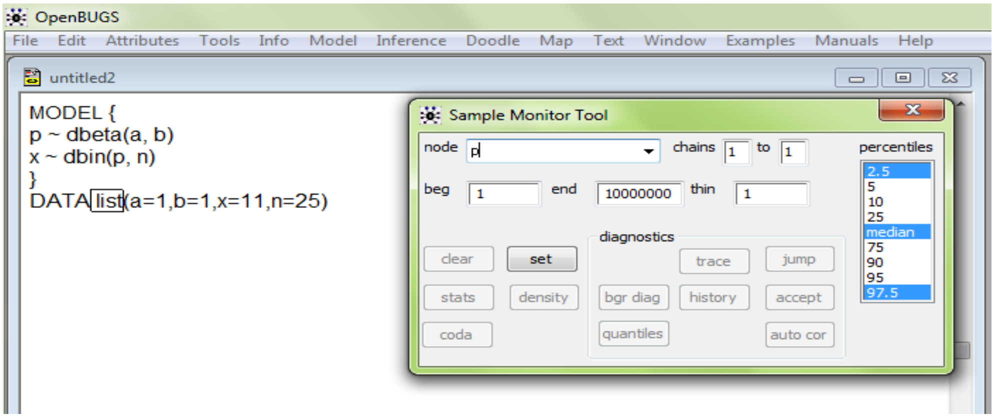


Y aparece la ventana Sample Monitor Tool.



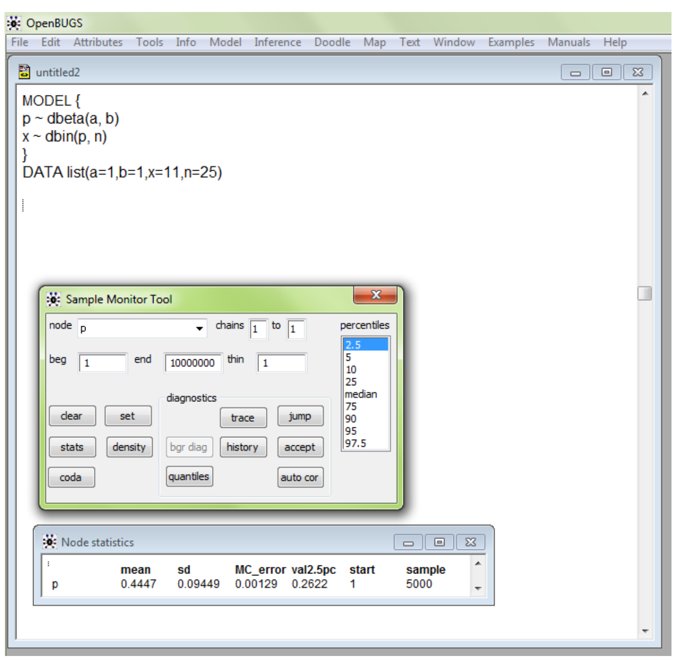
Tenemos que especificar los parámetros que deseamos simular, en principio todas las componentes de θ.

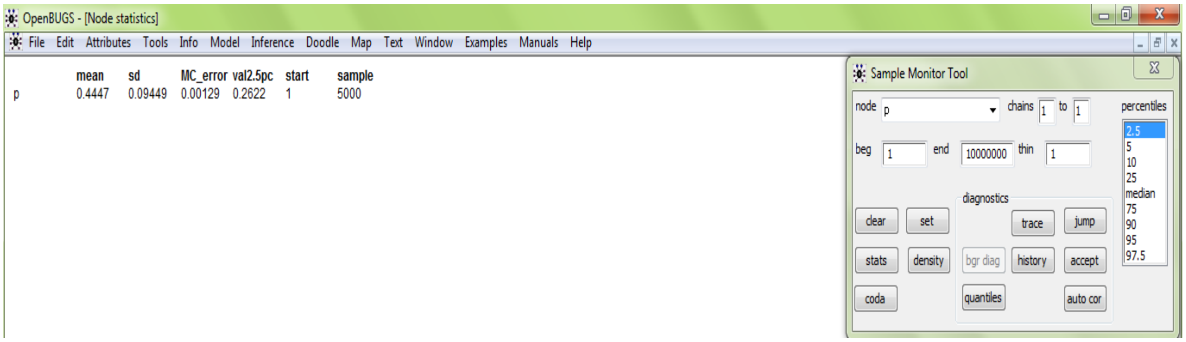
Escribir en el cuadro de "nodo" el parámetro que queremos monitorear en este caso es p y hacemos clic en set. y hacemos clic en set.



Por el procedimiento anterior, se especifica los parámetros para los que deseamos estimar sus distribuciones posteriores a través de los valores the MCMC generated. Los valores simulados de estos parámetros se almacenan ahora con el fin de producir un análisis detallado posterior.

Cuando tenemos un gran número de iteraciones (o simulaciones), volvemos a la ventana **sample Monitor Tool**. Y hacemos clic en **stats,** aparece un descriptivo de ladistribución posterior del parámetro que aparece en **node**, basado en las simulaciones creadas hasta el momento.





Con los datos obtenidos podremos interpretarlos así:

Samples Stats ("p")

Esto produce la salida

Estimación por intervalo al 95%. Se puede modificar los porcentajes que nos interesen.

Nombre de la variable

Desviación típica de la estimación puntual

Node mean sd MC_error val2.5pc median val97.5pc start sample

p 0.4447 0.09449 0.00129 0.2622 0.4437 0.6289 1 1000

Media: Estimación puntual de la variable

Error típico de la media de Monte Carlo $\frac{\sigma}{N^{1/2}}$

Valor de la mediana de las muestras posteriores.

Describiremos lo realizado en el ejemplo

Se observa un $x = 11$ son los éxitos de la muestra de tamaño $n = 25$. Al realizar estimaciones: clásica con una probabilidad de éxito de $11/25 = 0{,}44$. Como (Bayesiano) hemos supuesto que el parámetro p tiene una distribución a priori Beta $(1{,}1)$ (de parámetros α y β).

Se toma una solución algebraica para la distribución posterior (se tiene una distribución beta con parámetros supuestos con $\alpha = 12$ y $\beta = 15$), se podría realizar de modo manual y ayuda de los métodos de winBugs o OpenBUGS, se podrá ilustrar como una distribución podría producir más o menos la misma respuesta.

“Mean” el valor resultante representa la media de la distribución posterior. Las estimaciones con simulación sean 0.442
Manual, Tomando la formula (3) del ejemplo 8, se obtiene la posterior

$n = 25,\ \alpha = 12,\ \beta = 15,\ x = 11$

$$\left(\frac{25}{52}\right)(0{,}44) + \left(\frac{27}{52}\right)\left(\frac{12}{27}\right)$$

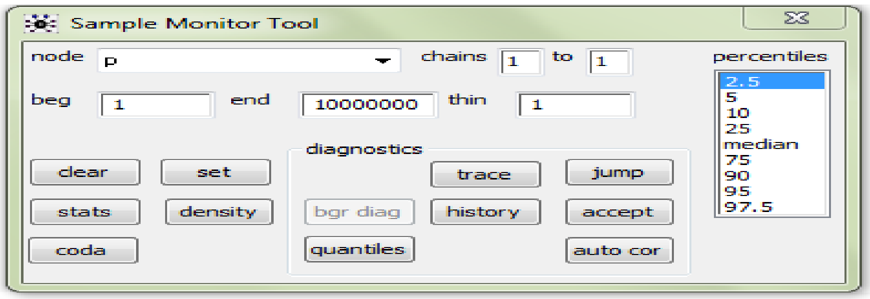
$$(0{,}48076923)(0{,}44) + (0{,}519230769)(0{,}4444)$$

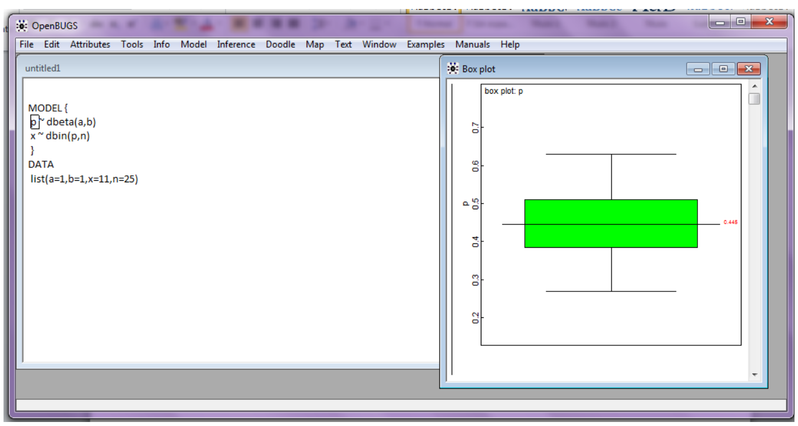
Para un total 0,442307. Si ha utilizado la solución algebraica, se llega a una media posterior, por formula (2) seria $\frac{12}{27} = 0{,}4444 \ldots..$
El valor de la etiqueta "SD" representa la desviación estándar de la distribución posterior la cual genera un valor de 0,09449.

El valor de la etiqueta "MC_error" es la incertidumbre asociada con esta simulación particular. Si este valor es mucho menor que la desviación estándar de la distribución posterior, entonces su simulación ha hecho bien. Si no es por lo menos un orden de magnitud más pequeña, entonces usted debería considerar el cambio de la simulación (por ejemplo, aumentando el número de muestras).

Los valores etiquetados "val2.5pc" y "val97.5pc" representan un intervalo que contiene 95% de la distribución posterior. Este intervalo se denomina intervalo de credibilidad (CRI) y representa la alternativa bayesiano para un intervalo de confianza. Los valores simulados 0.2634 y 0.6286 muestran que hay una buena cantidad de incertidumbre sobre la verdadera proporción p, lo que no es sorprendente dado el pequeño tamaño de la muestra. Estos valores para el CrI calculadas por la simulación se comparan muy bien con los valores exactos de la solución algebraica (0,2622 y 0,6289).

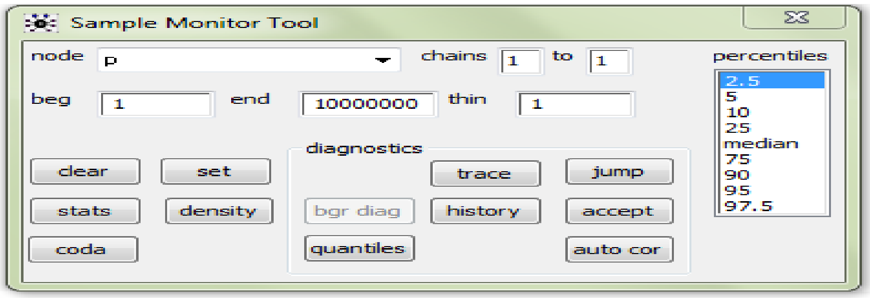
Los valores etiquetados "start" y "muestra" simplemente te recuerdan lo que se propuso para el comienzo y final de la simulación de la cadena Markvo.





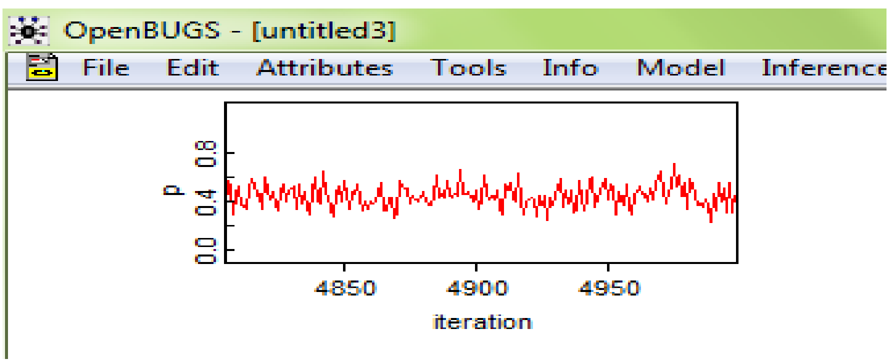
BOX PLOT

Es un gráfico que está fundamentado en cuartiles y mediante el cual se visualiza la distribución de un conjunto de datos.



TRACE (Dinamic trace)

Produce una parcela en la línea de los valores generados contra cada iteración número. Los puntos se trazan cada iteración k; donde k es el valor declarado en el botón de la herramienta de "refresh" "update" tool. Esta parcela también se conoce como dinámica trama traza en el manual WinBUGS

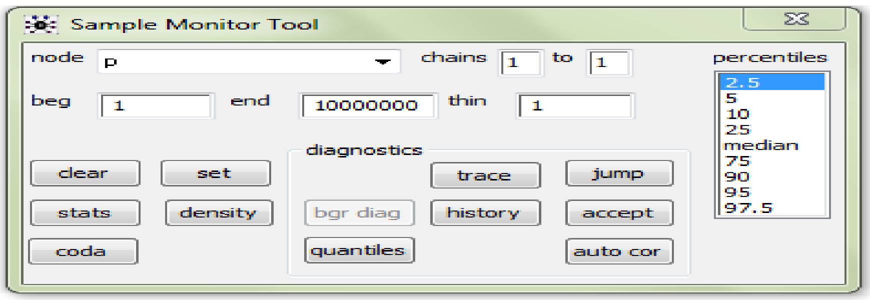


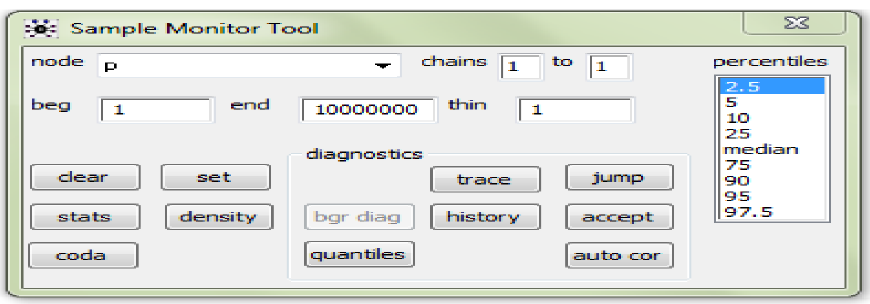
HISTORY

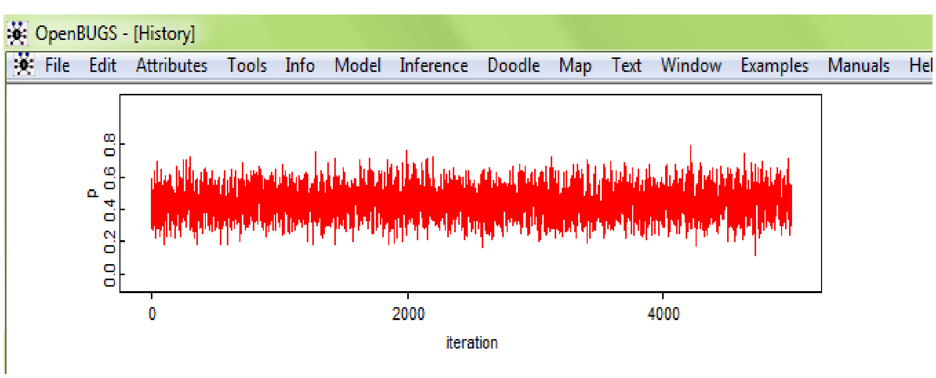
Mientras que la trama "rastro" proporciona una parcela en la línea de los valores generados, el botón de "historia" dibuja un trazo-trama llena de todos los valores almacenados

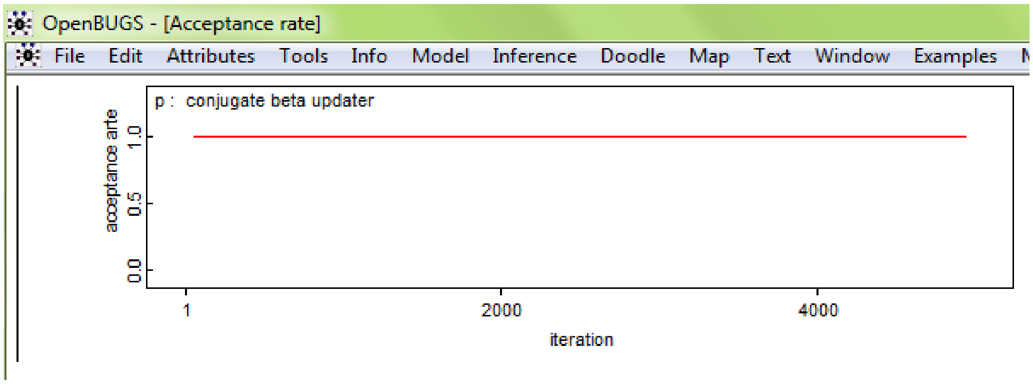
ACCEPT (Acceptan cerate)

Gráfico de la tasa de aceptación de la MCMC. Generalmente los modelos que utilizan un muestreador de Gibbs estarán siempre a 1 (100% de aceptación), mientras que otros métodos numéricos aceptan típicamente alrededor de 30 - 50% de los valores propuestos, una vez que ajustar su tamaño de paso. En general usted querrá para adelgazar la MCMC por lo menos 1 / acepta.



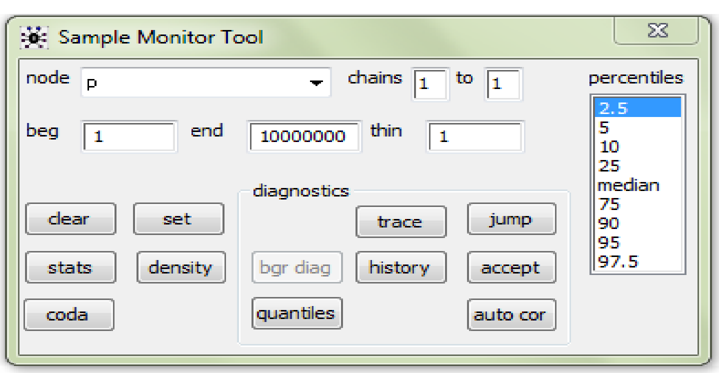






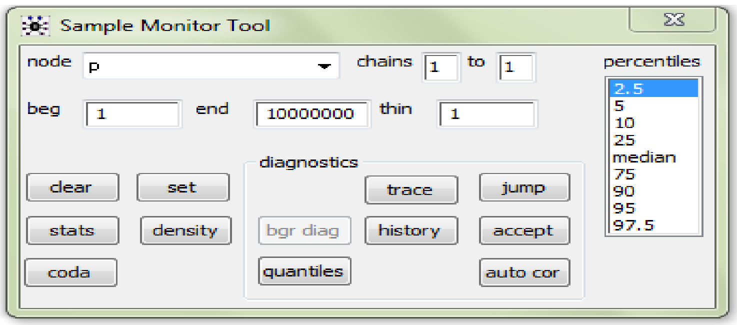
AUTO COR (auto correlación)

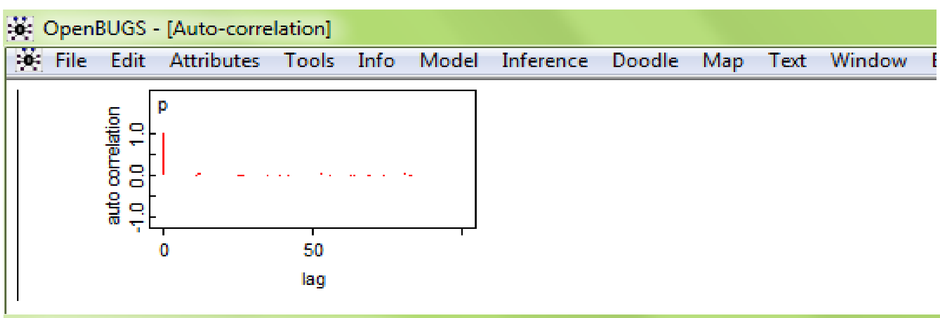
Esto genera un diagrama de autocorrelación para la MCMC, con el retraso en el eje x y el coeficiente de correlación en el eje Y. En el retardo 0 la cadena está siempre perfectamente correlacionada con la misma (autocorrelación de 1). Un desfase de un mostraría la correlación entre cada valor de la MCMC con el de al lado, un retraso de dos sería la correlación entre los valores que están separados por dos, etc. Buscamos cuando las asíntotas de autocorrelación a cero (típicamente el diagrama se ve más o menos exponencial). Usted tendrá que ajustar el "fino" en el Monitor Muestra Herramienta para el retraso en el que la correlación es casi cero (es decir, las muestras son independientes). Esta prueba suele ser más conservador que el procedente de la evaluación de la tasa de aceptación. Debido a que la "fina" afecta a todos los demás datos que querrá volver a calcular los valores que se están manteniendo (por ejemplo, gráficos de densidad o estadísticas de resumen) si cambia la delgada.

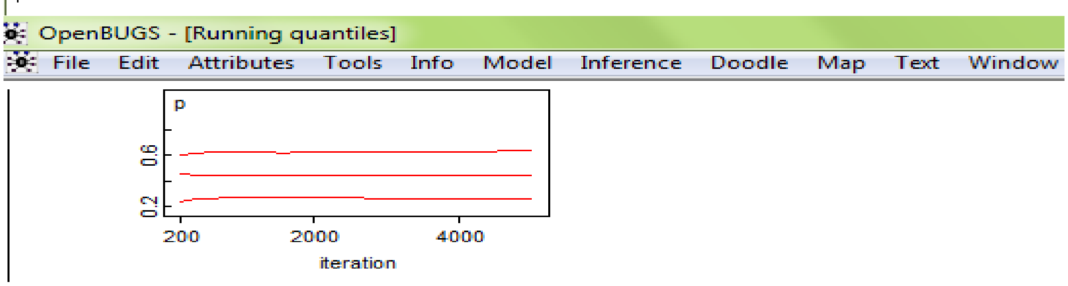


QUANTILES (Running quantiles)

Un gráfico de la evolución de la mediana y el 2,5% y 97,5% percentiles para cada iteración del algoritmo se obtiene por este botón.

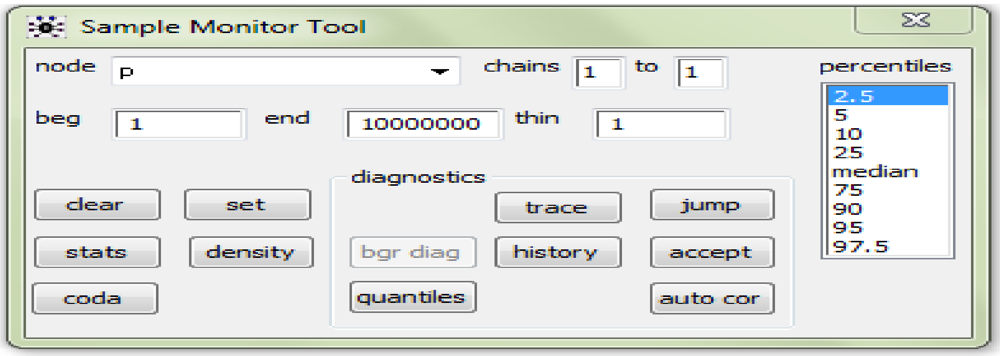


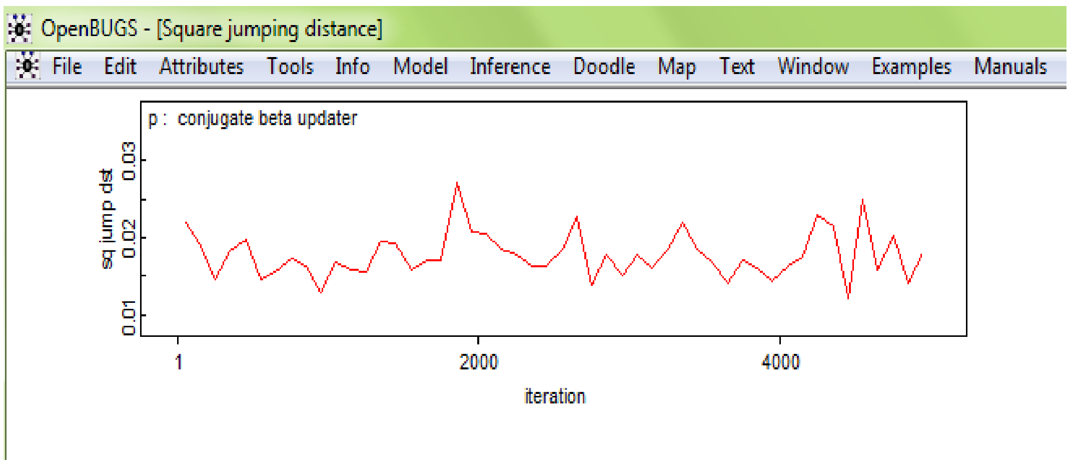




JUMP (Square jumping distance)

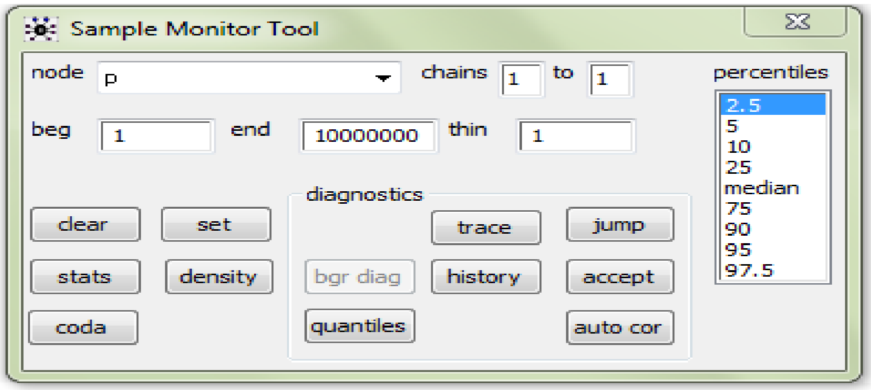
Crea una gráfica de la distancia de salto cuadrado del MCMC. Esto le permite ver si el modelo está dando pasos grandes o pequeñas.

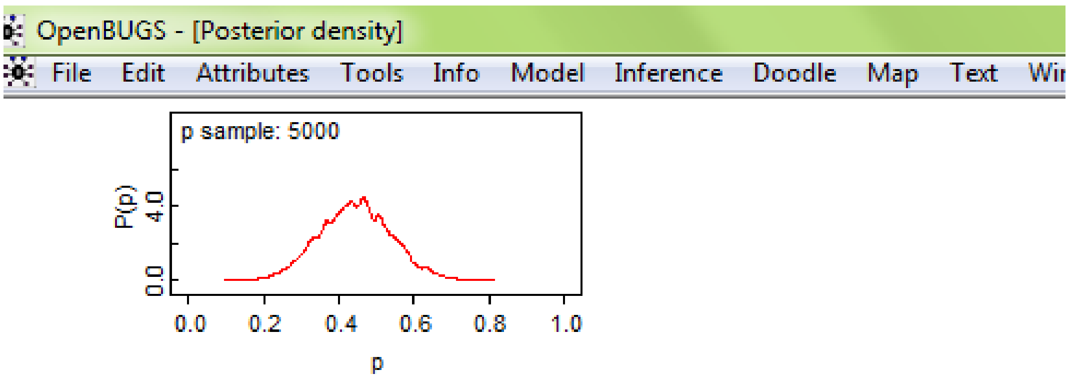




DENSITY (posterior density)

Este botón produce una estimación visual aproximada de la densidad posterior o función de probabilidad esto es mucho más a menudo un indicador de la falta de convergencia.


Sample Monitor Tool
node
p
chains
1
to
1
percentiles
beg
1
end
10000000
thin
1
2.5
5
10
25
median
75
90
95
97.5
diagnostics
clear
set
trace
jump
stats
density
bgr diag
history
accept
coda
quantiles
auto cor


OpenBUGS - [Posterior density]
File Edit Attributes Tools Info Model Inference Doodle Map Text Wir
p sample: 5000
P(p)
0.0
4.0
0.0
0.2
0.4
0.6
0.8
1.0
p

Los suscritos:

ANDREA JIMENA HERRERA GOMEZ	con C.C N°	**65.707.510 del Espinal**
	con C.C N°	
	con C.C N°	
	con C.C N°	
	con C.C N°	

Manifiesto (an) la voluntad de:

Autorizar [X]

No Autorizar [ ] **Motivo**: ______________________

La consulta en físico y la virtualización de **mi OBRA**, con el fin de incluirlo en el repositorio institucional de la Universidad del Tolima. Esta autorización se hace sin ánimo de lucro, con fines académicos y no implica una cesión de derechos patrimoniales de autor.

Manifestamos que se trata de una OBRA original y como de la autoría de LA OBRA y en relación a la misma, declara que la UNIVERSIDAD DEL TOLIMA, se encuentra, en todo caso, libre de todo tipo de responsabilidad, sea civil, administrativa o penal (incluido el reclamo por plagio).

Por su parte la UNIVERSIDAD DEL TOLIMA se compromete a imponer las medidas necesarias que garanticen la conservación y custodia de la obra tanto en espacios físico como virtual, ajustándose para dicho fin a las normas fijadas en el Reglamento de Propiedad Intelectual de la Universidad, en la Ley 23 de 1982 y demás normas concordantes.

La publicación de:

Trabajo de grado	X	Artículo		Proyecto de Investigación	
Libro		Parte de libro		Documento de conferencia	
Patente		Informe técnico			
Otro: (fotografía, mapa, radiografía, película, video, entre otros)					

Producto de la actividad académica/científica/cultural en la Universidad del Tolima, para que con fines académicos e investigativos, muestre al mundo la producción intelectual de la Universidad del Tolima. Con todo, en mi condición de autor me reservo los derechos morales de la obra antes citada con arreglo al artículo 30 de la Ley 23 de 1982. En concordancia suscribo este documento en el momento mismo que hago entrega del trabajo final a la Biblioteca Rafael Parga Cortes de la Universidad del Tolima.

De conformidad con lo establecido en la Ley 23 de 1982 en los artículos 30 ***"…Derechos Morales. El autor tendrá sobre su obra un derecho perpetuo, inalienable e irrenunciable"*** **y 37 "*…Es lícita la reproducción por cualquier medio, de una obra literaria o científica, ordenada u obtenida por el interesado en un solo ejemplar para su uso privado y sin fines de lucro*". El artículo 11 de la Decisión Andina 351 de 1993, "*los derechos morales sobre el trabajo son propiedad de los autores*" y en su artículo 61 de la Constitución Política de Colombia.**

- Identificación del documento:

Título completo: **ESTIMACION DE UNA PROPORCION BINOMIAL MEDIANTE METODOS BAYESIANO**

- Trabajo de grado presentado para optar al título de: **Profesional en Matemáticas con Énfasis en Estadística**

- Proyecto de Investigación correspondiente al Programa (No diligenciar si es opción de grado "Trabajo de Grado"): ________________

- Informe Técnico correspondiente al Programa (No diligenciar si es opción de grado "Trabajo de Grado"): ________________

- Artículo publicado en revista: ________________

- Capítulo publicado en libro: ________________

- Conferencia a la que se presentó: ________________

Quienes a continuación autentican con su firma la autorización para la digitalización e inclusión en el repositorio digital de la Universidad del Tolima, el:

Día: **29** Mes: **Junio** Año: **2017**

Autores: Firma

Nombre:	**ANDREA JIMENA HERRERA GOMEZ**	cc 65707510.	C.C.	**65.707.510 del Espinal**
Nombre:			C.C.	
Nombre:			C.C.	
Nombre:			C.C.	

El autor y/o autores certifican que conocen las derivadas jurídicas que se generan en aplicación de los principios del derecho de autor.